The representation and perception of visual motion: to integrate or not to integrate

by

James H. Hedges

A dissertation submitted in partial fulfillment

of the requirements for the degree of

Doctor of Philosophy

Center for Neural Science

New York University

September, 2009

Eero P. Simoncelli

J. Anthony Movshon

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

Acknowledgements

I would like to thank my advisors: Eero Simoncelli and Tony Movshon. I have long admired the incredible ingenuity and commitment that they bring to bear on answering difficult questions. I am grateful for having had the opportunity to work with them during my graduate studies.

I would like to thank my committee members: Nava Rubin, Mike Landy, John Rinzel and Lynne Kiorpes.

I would like to thank my examiners: Norma Graham and Bart Krekelberg.

I would like to thank my collaborators: Jenny Gartshteyn, Adam Kohn, Bill Newsome, Nicole Rust, Tim Saint, Mike Shadlen and Alan Stocker. Adam, Nicole, Tim, Mike and Bill contributed to the work described in chapter 1. Jenny collected psychophysical data related the work described in chapter 1. Tim helped develop the motion energy model described in chapter 1. Alan helped develop the Bayesian transparency model described in chapter 2.

I would like to thank the Center for Neural Science (CNS) class of 2003: Tarimotimi Awipi, Mitch Day, Yasmine El-Shamayleh and Riju Srimal.

I would like to thank some present and former members of the Laboratory for Computational Vision: Jose Acosta, Eizaburo Doi, Rob Dotson, Chaitu Ekanadham, Rosa Figueras, Jeremy Freeman, Deep Ganguli, Jose Antonio Guerrero-Colon, David Hammond, Yan Karklin, Misha Katkov, Siwei Lyu, Josh McDermott, Jonathan Pillow, Umesh Rajashekar, Martin Raphan, Jon Schlens, Brett Vintch and Rob Young.

I would like to thank some present and former members of the Visual Neuroscience Laboratory and Pete Lennie's lab: Neel Dhruv, Mike Gorman, Arnulf Graf, Mehrdad Jazayeri, Romesh Kumbhani, Matt Smith, Sach Sokol and Chris

Tailby.

I would like to thank some of the CNS/NYU faculty: Paul Glimcher, Mike Hawken, David Heeger, Sam Feldman, Pete Lennie, Larry Maloney, Dan Sanes, Mal Semple, Bob Shapley and Dan Tranchina.

I would like to thank some other people who are (or were) in, or around, CNS: Eric DeWitt, Anita Disney, Jeff Erlich, Andy Henrie, Chris Henry, Trent Jerde, Siddhartha Joshi, Brian Lau, Gabriel Lazaro-Munoz, Shani Offen, Hysell Oviedo, Lana Roisis, Robb Rutledge, Scott Schafer, Max Schiff, Abraham Schneider, Pascal Wallisch and Dajun Xing.

I would like to thank some present and former members of the CNS staff: Ken Anderson, Amala Ankolekar, Joey Azevedo, Krista Davies, Paul Fan, Erick Howard, Vic Keenan, Stu Greenstein, Joanne Rodriguez, Hillary Webb and Amy Yochum.

I would like to thank my family: Chuck and Carole Hedges; and Aug and Carolyn Firth. I would also like to thank Jerry and Mary Lou Barnes. I would like to thank two other New York-based neuroscientists: Annegret Falkner and James Herman.

There are many others who I will not name. Their contribution is, in any case, largely unquantifiable: girlfriends, friends, roommates, neighbors, musicians, painters, architects, designers, chefs, doormen, clerks, cab drivers, train conductors and janitors.

Abstract

I have addressed the physiological mechanisms for, and perceptual consequences of, integrating visual motion. Where possible, I have tried to determine the rules by which the visual system decides whether to integrate or not. My first set of experiments was motivated by the following observations. Humans and primates can see motions at small and large scales. Also, neurons in area MT have large receptive fields, which are known to play a role in the perception of visual motion. I conducted a series of electrophysiological experiments to determine whether MT neurons compute global motion, which was defined in terms of widely separated apparent motion. I used stimuli in which there could be opposing local and global motion. I found that MT neurons are unaffected by global motion, that their responses are entirely determined by local motion. My control experiments suggest that they do not compute global motion even in the absence of local motion. My second set of experiments concerned how the visual system decides whether to integrate or segment motions. I presented drifting square-wave plaids and asked subjects to indicate whether they appeared to move coherently, as a single object, or transparently, as two objects moving in different directions. I found that a plaid's component and pattern speed affected how it was perceived. Plaids were more transparent at faster pattern speeds and were coherent otherwise. I developed a Bayesian model that can explain these results. Key components of the model are based on preferences of the system to see slow and singular motion. My final set of questions was motivated by the idea that adaptation causes repulsion by reducing the gain of mechanisms that encode properties of a stimulus. In psychophysical experiments, I measured the pattern of biases in perceived direction that result from adapting to coherent and transparent drifting square-wave plaids. My results

suggest that adapting to plaids causes repulsion away from their pattern directions, even when they are not perceived.

Contents

Acknowledgements iii

Abstract v

List of Figures ix

Introduction 1

0.1 Basic observations 1

0.2 Three sets of questions and three sets of results 17

1 MT computes local but not global motion **20**

1.1 Introduction 21

1.2 Methods 23

1.3 Results 31

1.4 Discussion 45

2 Speed-dependent plaid perception, the interplay of two preferences **51**

2.1 Introduction 52

2.2 Methods 59

2.3 Results 70

2.4 Discussion 83

3 Adaptation-induced repulsion is away from pattern directions 100

3.1 Introduction 101

3.2 Methods 104

3.3 Results 111

3.4 Discussion 121

4 Conclusions and future work 128

4.1 The scale of visual motion 128

4.2 To integrate or to segment 130

4.3 Adapting to visual motion 132

Bibliography 134

List of Figures

1.1 Do MT neurons represent global motion 22

1.2 An apparent motion stimulus with local and global motion 26

1.3 Responses of an MT neuron to local motion with different durations and to local-global motion with different global speeds 32

1.4 Quantifying the scale of MT's directionality 35

1.5 The effects of global motion on the responses of a population of MT neurons 36

1.6 Responses of an MT neuron for experiments in which the stimuli contained no local motion 38

1.7 MT's directionality in the absence of local motion 40

1.8 Local-global stimuli in the frequency domain 42

1.9 Comparison of the scale of directionality predicted from a motion energy model to actual data 43

2.1 The perception of a drifting square-wave plaid 53

2.2 Psychophysical protocol and example plot of plaid percepts in component versus pattern speed 56

2.3 The effects of a plaid's speeds on how it appears to move 71

2.4 Bayesian observer model for plaid perception 73

2.5 The effects of manipulations of the two prior distributions 77
2.6 Fits of Bayesian and alternative models to data 79
2.7 Comparison of model performance 82

3.1 Hypotheses for the perceptual consequences of adapting to a square-wave plaid 103
3.2 Method of adjustment task for measuring adaptation-induced biases in perceived direction of motion 107
3.3 Selecting coherent and transparent plaid adaptors 110
3.4 Circle-shaped plot for showing adaptation-induced biases in perceived direction of motion 112
3.5 Adaptation-induced biases from a coherent plaid 113
3.6 Adaptation-induced biases from a transparent plaid 117
3.7 A comparison of the performance of component and pattern predictions for coherent and transparent plaids 118
3.8 Adaptation-induced biases from transparent random dots 120

Introduction

This thesis is presented as three self-contained papers. They are not part of a single line of study, but are connected in that they address aspects of the representation and perception of visual motion. My focus has been the perceptual and physiological consequences of integrating visual motion information. Where possible, I have tried make clear the rules by which the system decides whether to integrate or not. I have also tried to identify the mechanisms which are associated with the perceptual phenomena in which I was interested.

In this introduction, I briefly state what visual motion is, how representations of it are formed, and what representations of it convey to an organism. I summarize sets of observations on how it is represented and perceived. At the end of this chapter, I link the questions that I address in the chapters that follow to the background I present here and I briefly summarize my results. In many cases, I have omitted some alternative views of the background I describe. I likewise have not included all known pieces of evidence, for or against, the conclusions I discuss. My aim is not to provide a comprehensive summary of these issues, but to lay out the threads of knowledge that are foundational to the questions that I explored.

0.1 Basic observations

In simple terms, visual motion may be defined as changes in the position of light over time. To represent these changes, a system can first transduce the brightness within many different small cone-shaped regions in the surrounding three-dimensional world. The retina does this by estimating the amount of light that lands on its surface at many different positions in a given amount of time.

When there is motion, such as when an object moves in a scene, the pattern of the projected retinal image is shifted relative to the previous image that it formed. It is the job of higher areas to estimate the motions in a scene from these changes in the retinal representations. I should point out that this is true for luminance-defined motion, but visual motion can be defined in other ways, an issue to which I will return. Changes in luminance are sufficient, but not necessary for visual motion. They may or may not evoke a sense of motion when the system has access to them and the sense of motion can result without them.

A representation of motion provides a wealth of information to an organism. There is ample evidence, for example, that humans use information about motion to interpret the world. Among the functions it serves are: sensing the real motion of objects; sensing the depth and relative distances to points in the environment; sensing self motion, such as when walking in an environment; estimating the time to collision of visual targets; segregating different objects in a scene; distinguishing figure from ground; and driving eye movements. For many organisms, these functions are essential for survival. In a sense this must be the case, since the mechanisms that represent motion consume considerable resources, even though they have been optimized to be as efficient as possible.

0.1.1 The aperture problem and a 'solution'

The 'aperture problem' is an inherent ambiguity that results when measurements of local motions are made [1, 2, 3, 4]. The solution to this ambiguity shapes the form and processes of the system that represents motion. Consider an extended contour, such as a line segment, moving within an aperture. There is not enough information to determine the true motion of the surface of which the contour is

part. In other words, there are many different physical motions that can lead to the same physical stimulus. The key point is that any motion parallel to a one dimensional (1D) pattern is invisible and, therefore, only motion normal to its orientation can be detected.

The visual system faces this problem, since it makes local measurements (i.e., through an aperture), and many structures in the world are 1D when measured locally. But it can, and does, determine the true motion of objects when more than one unique 1D motion exists. The basic idea is to pool 1D motions that are consistent with the same pattern of motion, as would result from a single translating object. Since these estimates can be assumed to relate to each other in a specific way a unique solution emerges. It is, nonetheless, not clear which signals to pool. The idea is to pool the ones that are consistent with pattern motion, although the pattern motion is unknown without pooling the local motions [5, 6].

It is easiest to consider the solution in the context of a few classic visual patterns. The set of 1D stimuli to consider includes extended gratings, edges and bars. None of these are truly one-dimensional in the real world, which is to say their extent is not infinite, but they are essentially so, if they extend beyond the edge of the region of visual space that is represented or perceived. The set of two dimensional (2D) stimuli to consider includes plaids, which are the superposition of two component gratings in overlapping visual space, and random dots, which can be thought of as the superposition of a set of many overlapped drifting gratings.

In velocity space, each of these moving stimuli has a corresponding vector or pattern of vectors, whose length relates to the stimulus' speed and whose orientation relates to its direction [7]. A drifting grating, for example, maps to a line of possible velocities, a constraint line. The constraint line is parallel to the stimulus'

orientation and orthogonal to the vector representing its primary motion. For a plaid there are two such constraint lines, one for each component, and they intersect at a unique point. By integrating the two local motions the intersection of constraints (IOC) solution can be determined. This corresponds to the true motion of the plaid [3, 8, 7, 9]. Plaids are a simple case that illustrates the IOC solution, but IOC is not limited to them. The point is that it provides a solution for more complicated motions.

But IOC is not the only way to combine motion vectors, and some other options include summing or averaging them. For some stimuli, there may be no difference between the solutions provided by IOC and these other rules. One case for which the solutions are different is a 'type II' plaid. This is a plaid for which the direction in which the constraint lines intersect is outside of the narrowest sector of the directions of the two component motions. There is some evidence that the perceived direction of motion (DoM) of type II plaids is biased towards the vector average direction [10, 11]. Results like these remind us that the details of which rule is most similar to what is perceived in different circumstances, have not been worked out. That said, IOC is a good approximation for many stimuli.

A two-stage model for motion perception emerges from these observations about integrating motions. The system first estimates local motions and then combines them [3, 7, 9]. The first step is the decomposition of the image into 1D spatial components of varying orientations. The speeds in the direction orthogonal to the component orientations are computed. The directions and speeds of many different components at a given spatial location are recombined to find the IOC solution.

This framework for estimating motion is formal and does not connect to any particular physiological implementation. But subsequent studies have established

that the system that represents motion does so in a way that parallels the IOC solution. I will return to the evidence that supports this in some subsections that follow, where I review the physiological and theoretical support for the two-stage model. Before turning to that, I review what is known about deciding whether to integrate or segment motions.

0.1.2 Integration versus segmentation

There is a need for integration, but the system cannot integrate all of the motions it represents. Motion signals may come from different objects, for example, in which case combining their motions into one would be incorrect. The choice of whether and how to integrate motions by the system is related to the rules by which the integration is performed. Answering this question tells us something about the properties of the mechanisms that are responsible for the integration. This issue has been explored extensively with moving plaids, which have percepts that are consistent with integration or segmentation. They appear either as a coherent pattern moving with a singular velocity or as transparently overlaid patterns moving in different directions. A simple approach used in many psychophysical studies is to present plaids with different properties and ask subjects to indicate how they appear to move.

I should mention at the outset that synthesizing the results of many different psychophysical experiments on integration and segmentation is not straightforward. There are many differences between them, some of which are likely to influence their outcomes, sometimes in unknown ways. The differences include the plaids themselves, which tend to be of different sizes, eccentricities and contrasts. Plaids are usually presented for different amounts of time or as part of different

tasks. In terms of tasks, subjects may be instructed to report different things, such as a plaid's pattern direction relative to a reference direction, instead of whether it looks coherent or transparent. In designs where subjects report which of the two percepts they saw, it is assumed that there are only two possible percepts. Taken together, these issues make it difficult to trust simple explanations for the effect of a given parameter. Increased contrast may in some cases make plaids more coherent, for example, but that might only be true within a limited portion of the overall stimulus space, or true for sine-wave plaids, but not square-wave ones. My view is that these issues should inspire the broadest possible sampling of any parameter space that is explored.

One early study on plaid integration showed that the relative contrast of the components has an effect [9]. For a plaid with one component of contrast 0.3, increasing the contrast of the other component increased coherence. It is not clear whether this was a consequence of the overall contrast level, or a consequence of the similarity of the components' contrasts, or some mixture of the two. A similar effect for the relative spatial frequencies of the components was described. The coherence decreases as the difference in spatial frequencies increases. Not much data was shown in support of these results initially, although subsequent studies have examined the effects of relative contrast and spatial frequency in greater detail.

Many of the studies that followed continued to explore relative effects, where the components differed in ways other than having different directions of motion. This may have to do in part with the difficulty of evoking transparent percepts with sine-wave plaids when the components are not different. Other types of plaids, such as square-wave plaids, seem to be more transparent, so there is a greater chance

of observing both percepts.

Smith examined the effects of relative contrast in later studies and found that they were consistent with the earlier observations: coherence was maximal when the components had the same contrast and decreased as the difference between their contrasts increased [12]. The effect of relative spatial frequency was that the maximum spatial frequency difference at which plaids were seen as coherent was up to three or four octaves [13]. The range of spatial frequency differences for which coherence remained, decreased as contrast decreased or component speed increased. It seems like part of Smith's motivation to examine relative spatial frequency was to challenge the earlier suggestion by Adelson and Movshon that component motions are combined in spatial frequency tuned channels, prior to the formation of an estimate of the plaid's pattern motion [7]. Smith's results challenge this idea since they suggest that plaids can appear to move coherently even when the spatial frequencies of the components are very different. This suggests instead that there is pooling across broad ranges of spatial frequencies before pattern motion is computed.

Interestingly, most of these early studies on plaid integration assume that if a plaid with certain parameters does not evoke a coherent percept then the pattern motion is not being represented internally. Another way to say this is that they assume that if a representation of pattern motion is formed then the coherent percept occurs.

The effect of plaid angle has also been measured for plaids with components that have different contrasts, spatial frequencies or speeds [14]. Angle had a significant effect. Plaids were coherent for angles smaller than about 45 deg and became increasingly transparent as angle increased. There were minimal differences in the

effect of angle for contrast and spatial frequency differences up to 9:1. This was taken as evidence that the angle between a plaid's components' DoM is the primary determinant of the percept. The effects of angle were not connected with those of pattern speed, which covaries with angle. Also, only a few component speeds were used. The influences of angle may have been different at other speeds. Data from plaids in which the two components have the same parameters were not included, although that seems like an important control.

Another aspect of a square-wave plaid that influences how it is perceived to move is the luminance of the intersections of its components [15, 16, 17, 18]. When the luminance of the intersections of drifting square-wave plaids is chosen so that they look transparent (as stationary patterns) the coherent percept decreases. Stoner and colleagues speculated that this means that the visual system has access to tacit knowledge of the physics of transparency and that it uses this knowledge to segment a scene. They did not explain, or have not explained, how the system would store or develop such knowledge. Nor have they described the mechanisms by which the knowledge would influence which percept occurs. Moreover, the effect of the luminance of the intersections is highly speed dependent. If you slow down the plaid, the effect starts to disappear (plaid becomes strongly coherent). Therefore, the explanation in terms of luminance combinations that are consistent with transparency is, at the very least, only partially correct.

The view advanced by Stoner and colleagues about the effects of luminance on plaid perception evolved from the rules of transparency idea, to similar ones about depth ordering and disparity [19]. Depth segregation and transparency increase when a higher contrast component is presented so that it appears to be in front of a lower contrast component. As is the case for many studies on plaid perception,

the effects of luminance or depth were related to, or put in the context of, other factors that influence how a plaid is perceived. In many cases, the other studies are not even mentioned.

There have been a couple of reports by Farid and colleagues about the effects of component and pattern speed in square-wave plaids [20, 21]. The plaids they studied had matched components. They found that plaids are more transparent when their pattern speeds are faster than a cutoff speed (about 5 deg/s), and their component speeds are slower than the cutoff speed [20]. Plaids with pattern speeds less than a cutoff speed were coherent. Plaids with component speeds beyond a cutoff could not be perceived clearly, which may have meant that they were neither coherent nor transparent. They suggested that the luminance of a plaid's intersections only affected the percept when the pattern speed was near a cutoff speed.

Farid's results conflict with the notion that angle is the primary determinant of plaid perception. They suggest instead that the plaid's speeds are of greater importance. That the pattern of results was consistent with there being a cutoff speed means that perception varies in a way other than by angle. They also explored the effects of overall contrast and of the period of the components [21]. Their results in the followup study also suggested that there is a cutoff speed, which they found depended on a plaid's contrast and period. They found that plaids with higher contrasts or which have a smaller period are more coherent.

0.1.3 MT neurons have a central role in motion processing

MT is a primary component of the physiological pathway for motion processing. It was discovered concurrently by two groups [22, 23]. One was Dubner and

Zeki, who were the first to report that MT neurons are direction-selective. They suggested that MT has a columnar organization for direction and that it plays a role in guiding pursuit eye movements [22]. These suggestions were supported by later studies [24, 25]. The other group was Allman and Kaas, who reported that MT neurons respond more to drifting bars than to flashed spots [23].

A number of more recent studies confirmed that MT is highly specialized for representing visual motion. Some of the evidence in support of this includes that MT has a high concentration of direction-selective neurons in New and Old World Monkeys [26, 27, 28, 29, 30, 31, 32]. A current view is that MT responses are determined not only by direction, but by these additional properties of a stimulus: retinal position; speed of motion; binocular disparity; and size [33, 34, 35].

An important connection between the two-stage model and the physiological mechanisms that implement it, came from electrophysiological results that showed that some cells in MT are tuned for component directions of motion, whereas others are tuned for the pattern direction of motion [9, 36, 37]. There are neurons, in other words, that analyze 1D motion, as in the first stage, and neurons that analyze 2D motion, as in the second stage. Component direction selectivity is like orientation selectivity with direction selectivity. Component cells respond to 1D motions that are part of a complicated 2D pattern of motion in the same way as they do when the 1D motions are presented in isolation. Pattern direction selectivity is like component direction selectivity, but for 2D motion. Pattern cells respond to 2D motions in the same way as they do for 1D motions.

An important advance made by Movshon and colleagues, was to construct hypothetical predictions from data for either type of tuning for each cell, which allowed them to classify cells as component or pattern tuned. The pattern of

responses to gratings moving in different directions would be the same as for a plaid moving in different directions for a pattern cell. In contrast, the pattern of responses of a component cell would be like those to drifting gratings, but would be shifted with respect to the pattern direction so that the components would be moving in the component cell's preferred direction. Using these predictions all V1 cells were shown to be component cells, whereas some MT cells were shown to be pattern cells [9].

An important followup study tested whether the V1 neurons that project to MT are themselves direction-selective and whether they are component or pattern cells [38]. They found that the V1 projection neurons are strongly direction-selective and are component-tuned. They were typically 'special complex' and they generally responded to a broad range of spatial and temporal frequencies. This is consistent with the two-stage model. Similar work suggests that MT's projection neurons are tuned to speed and to binocular disparity [39, 40].

It has recently been asked how the responses of MT neurons differ from their inputs, since they are largely tuned for the same properties. Another way to look at it is to ask what MT adds.

One possibility is that MT computes motion over larger spatial scales. The plausibility for this is increased by the observation that the receptive fields (RFs) in MT are much larger than in V1. The standard ratio is about tenfold larger in linear dimensions [41]. It has also shown that MT neurons are directional at larger spatial separations than V1 neurons. The separations of MT neurons were, nonetheless, more consistent with what has been termed short-range motion. And a more recent set of measurements suggests that they actually have similar upper limits for the spatial separation at which they are directional [42]. The short-range

motion to which I referred, comes from evidence that humans have at least two different motion processing mechanisms, which operate at different spatial scales and have different characteristics [43, 44, 45].

Some unpublished results by Shadlen and colleagues tested the possibility that besides inheriting a representation of short-range motion from V1, MT neurons form, within themselves, a representation of long-range motion [46]. Their results suggested that they do not represent long-range motion, an issue I explored.

The responses of MT neurons are critical to the perception of motion. Small excitotoxic lesions of MT, for example, substantially elevate motion detection thresholds [47]. Also, the motion sensitivity of MT neurons is sufficient to explain the sensitivity of behaving monkeys [48, 49]. Results suggest that there is a statistical association between MT responses to identical stimulus presentations and the behavioral choices that monkeys make on different trials [50]. Simulations of these results suggest that responses from MT cells can be pooled to make perceptual judgements of motion stimuli [51]. And modifying the activity of MT neurons via microstimulation influences judgements of motion [52, 53, 54]. Injecting current in MT is akin to adding certainty about the direction of a moving stimulus.

0.1.4 Estimating velocity in theory

The basis of many models of estimating motion is the observation that for any moving rigid object the spatiotemporal frequency of all local motion estimates must lie on a plane in frequency space [55]. In this context, V1 simple cells are like space-time filters [56, 57]. Complex cells are similar in terms of being like filters, but they include phase insensitivity [58]. So V1 neurons measure the amount of motion energy in their passband [59]. They are incapable of representing the true

pattern motion since they can only see that part which falls within their frequency band.

In the model, an MT pattern cell can be formed by summing the responses of V1 cells over a plane in frequency space [60, 61, 62]. This planar summation is a way of implementing IOC. This model for MT cells can predict their responses to plaids and random dot stimuli with varying levels of coherence. The model makes some predictions about how MT cells should respond to different stimuli. One is that component cells should have bimodal responses to dots moving at fast speeds. Also, pattern cells should have bimodal responses to a grating moving at slow speeds. These predictions have been confirmed by physiological measurements of cells' responses [63].

Psychophysical evidence that supports mechanisms that sum energy over a plane includes detection experiments done with dynamic random noise patterns [64]. Results with these patterns suggest that there is a static mechanism in which energy is summed over an entire plane even when it is sub-optimal to do so. Subthreshold summation of signal-contrast energy improves when the energy is distributed broadly across many different orientations. Detection did not improve when the energy extended into non-planar regions. In other words, detection improves as more orientations are added to a moving pattern so long as they are consistent with the overall velocity of the pattern.

0.1.5 Adapting to visual motion

There have been many studies on the effects of adaptation, which have shown that it can cause repulsive shifts in the perception of different properties of a stimulus. These properties include: spatial frequency [65, 66]; size [67]; direction

[68, 69]; orientation [70, 71, 72, 73, 74]; and contrast [75]. Similar studies have been done with moving adaptors. They have shown that adapting to motion can cause: illusory motion of static tests ('Waterfall Illusion') [76, 77]; illusory motion of dynamic random dots [78]; and biases in perceived velocity or speed [79, 80, 81, 82, 83, 84, 85, 86]. There have also been a few studies that have examined the effects of adapting stimuli that appear to move transparently. They have used the following as adaptors: transparent random pixel arrays [87, 88]; transparent random dots [89]; and transparent plaids [90]. One finding from these studies is that the effect of adapting to transparent stimuli is unidirectional and in the direction opposite of the vector sum of the motions in the stimulus. But this issue has not been examined with respect to the direction of drifting tests. And the relationship between adaptation effects on static, or zero speed, tests with those on drifting tests is not known.

Most of the above-mentioned studies examined the effects of motion adaptation separately, either in terms of the perceived DoM, temporal frequency or speed. But Schrater and Simoncelli offer some insight into the perceptual consequences in 2D velocity space of adapting to moving visual patterns [83]. They evaluated three hypotheses: that adaptation effects are the result of sensitivity changes in spatiotemporal frequency tuned mechanisms; that adaptation occurs in mechanisms that encode direction and speed independently; or that it occurs in mechanisms that encode 2D velocity. Their approach was to probe the representation of velocity by measuring the pattern of perceptual shifts.

Their results show that adaptation causes shifts in the perceived direction of plaids that are relatively independent of the spatial pattern of the adaptor [83]. The shift in the perceived DoM of a plaid after adapting to a grating is away from

the direction of the overall plaid pattern and not away from a plaid's components. In 2D velocity space, the shifts in perceived velocity radiate away from the adapted velocity. The shifts are inseparable in speed and direction. These results are most consistent with the third hypothesis they proposed, that adaptation occurs in mechanisms that encode 2D pattern velocities.

Their experiments were based on a few simple assumptions, following those that were expressed by Blakemore and colleagues [65]. One is that the visual system represents a stimulus parameter using a population of neurons tuned for that parameter. The perceived value of the parameter is determined by the relative responses of those mechanisms within the population. Another assumption is that prolonged exposure to a stimulus reduces the responsivity of those mechanisms that represent it by an amount that is a monotonic function of their sensitivity to it. Given these assumptions, the pattern of adaptation-induced shifts can be taken as a representation of the parameter that is affected [67, 66, 65, 91].

Earlier work by Blakemore and colleagues made this connection. The perceived spatial frequency of a sine-wave grating is shifted away from the frequency of an adapting grating. They inferred from this that humans have mechanisms tuned for spatial frequency. Similar work showed that DoM repulsion occurs with drifting dots as adaptors and tests [68]. Also, repulsive speed effects had been shown [92]. Separate results by Thompson and Smith suggested that adaptation always causes a decrease in the perceived speed of tests [93, 80]. Recent results by Stocker and Simoncelli show that speed adaptation effects can be explained as a superposition of two repulsive patterns, implying the adaptation of two underlying mechanisms [86]. Other results by Smith and Edgar were consistent with repulsion for some combinations of adaptor and test [81].

0.1.6 Bayesian speed perception

Several Bayesian models for motion perception have been developed [94, 95, 96, 97, 98, 99, 100]. The approach of these models is to create an optimal Bayesian estimator, also called an 'ideal observer'. The models are usually tested against perceptual data gathered on simple stimuli (with a single motion), although the models themselves generalize. Some of them can make predictions for rotation and other smooth deformations, although not for occlusion boundaries or transparency. Most of them make the assumption that all changes in the intensity of the image over time are the result of smooth translational motion. The most important assumptions, though, are that local image measurements are noisy and that image speeds are usually slow. These last assumptions are represented in terms of probability distributions. Bayes' rule allows one to determine the ideal observer.

A recent Bayesian model for 2D motion estimation can account for a number of perceptual phenomena [100]. The Bayesian model captures the influence of contrast on perceived speed. The result is that lower-contrast patterns appear to move more slowly than higher-contrast patterns [93, 101]. The Bayesian model also captures the influence of contrast on the perceived direction of drifting plaids [102]. The result is that the perceived direction of a plaid can be biased towards the component with higher contrast when they are not the same. Finally, the Bayesian model captures the biases in the perceived direction of type II plaids towards the vector average solution [103].

A different Bayesian model accurately predicts human speed perception [104]. Stocker and Simoncelli estimated the shape of the prior for speed as well as the contrast and speed dependence of the likelihood function from forced-choice psychophysical experiments. The key to estimating these functions was to relate the

shape of their distributions to the psychometric function, which is to say to the variability across trials for different levels of the stimulus. Their results suggest that the width of the likelihood function, which relates measurements of the moving stimulus to estimates of its speed, is proportional to the log of speed and falls monotonically with contrast. The prior distribution falls off with speed as a power law, with the exception that the slope becomes shallower at the slowest and (for some subjects) fastest speeds.

0.2 Three sets of questions and three sets of results

I was interested in the scale of motion that is represented by MT neurons. Following the above-mentioned reasoning, I knew that humans and monkeys can see widely-separated apparent motion, which I took as evidence that there is a neural representation of widely spaced motions. MT neurons are a candidate for that representation since they have large RFs and may be able to respond selectively to motion over larger spatial separations as compared to neurons in upstream areas.

I tested the hypothesis that MT neurons not only inherit a representation of motion at fine spatial scales, but that they compute a new motion signal that depends on the direction of long-range motion. I used stimuli in which the directions of local and global motion could be pitted against each other, meaning they could be independently presented in the preferred or null direction of a cell. I made electrophysiological recordings from MT neurons to local-global stimuli with many different global speeds. My results suggest that MT neurons respond selectively

only to local motion and are unaffected by global motion.

I have also addressed the problem of deciding whether to integrate or segregate motions. In this introduction, I reviewed the effects of changes in different parameters on how plaids appear to move. Like many of those studies I did a series of psychophysical experiments in which subjects indicated their percept of a moving plaid. I measured their responses for a range of component and pattern speeds. My basis for exploring these parameters was the earlier work by Farid and Simoncelli, although my results were quite different.

I found that plaids are transparent when their pattern speed is slightly faster than their component speed, when their component speed is relatively slow. Plaids with other combinations of component and pattern speed among those that I explored were coherent. I augmented a previous Bayesian model for speed perception to create a model forming integrated or segmented plaid percepts. My model can account for my results. The model was based on the assumption that measurements of the component motions of a plaid are noisy, and that the human visual system, which represents them, imparts preferences for slow and singular percepts. Simulated percepts from the model suggest that subjects' prior distributions for speed and the values of their prior for seeing singular motions can explain their pattern of percepts across trials.

The final set of questions that I addressed deals with the perceptual biases that result from adapting to coherent or transparent plaids. The approach that I used is similar to the one described by Schrater and Simoncelli [83]. I tested whether the pattern of repulsion would depend on how the adaptor was perceived. I also considered a second possibility, that the pattern of repulsion would be away from the two component directions of motion, regardless of how the plaid appears. I

assumed that there are mechanisms that estimate a plaid's component and pattern velocities, which can be thought of as component and pattern cells.

My results suggest that adapting to coherent and transparent plaids causes repulsion away from their pattern directions. There may be some slight differences between the directional biases generated by coherent and transparent plaids. But the locus of the adaptation is always consistent with the pattern direction. My analysis suggests that the pattern of shifts can be explained as an additive mixture of component and pattern adaptations. Control experiments with transparently moving random dots suggest that repulsive biases in perceived direction can be away from more than one direction, but that they are always away from the directions represented by pattern-tuned mechanisms.

Chapter 1

MT computes local but not global motion

MT neurons receive direction-selective input from V1 and other extrastriate areas. This input may give rise to MT's directionality for motions with small spatial separations, but not for motions with large separations. I assessed the relative contributions of local and global motion on the directionality of MT neurons in anesthetized macaques. I presented brief narrow ovular 'blips' of drifting grating at sequential locations along the preferred-null axis of the RF. When the spatiotemporal frequency in each blip matched those preferred by a cell, responses were entirely local. Global motion did not elicit directionality by itself or when its separations were beyond approximately 1 deg. A simple model based on the convolution of a stimulus' motion energy with a cell's sensitivity profile, predicts these responses. These results suggest that directionality in MT is inherited from its inputs. Global motion must, therefore, be represented in some other visual area.

1.1 Introduction

Sensory computations are executed incrementally by the sequential convergence and divergence of populations of neurons, each of which plays a role in the broader processing within a sensory modality. Advancing through subsequent stages of convergent and divergent input and output is generally accompanied by increases in the complexity of sensory representations. In vision, for example, neurons in the LGN respond preferentially to small circular spots of light, whereas downstream neurons in V1 respond better to drifting bars or sinusoidal gratings [105, 106]. Transformations in visual response properties continue through extrastriate areas. Neurons in V4 and IT respond better to shapes, like those of objects or faces [107, 108, 109, 110]. These transformations have been modeled as cascades of linear and nonlinear processing steps [111].

As I described in the introduction, MT is central to the pathway for processing and perceiving visual motion [9, 47, 52, 49, 112]. A simplified model of the formation of MT's processing is that it results from a two-stage process [9]. Direction selectivity is established in V1, where approximately one quarter of the neurons respond more strongly to motion in one direction and minimally to motion in the opposite direction [106, 113]. V1 neurons project, both directly and through other areas, to MT, where the majority of neurons are direction-selective [22, 26, 29, 114, 38].

One of the differences between V1 and MT processing is that some MT neurons are tuned to the pattern direction of a plaid, whereas V1 neurons are not [9, 115]. Pattern direction selectivity is thought to be formed by a weighted combination of neurons tuned to the component motion. In other words, it results from integration over motion direction [62].

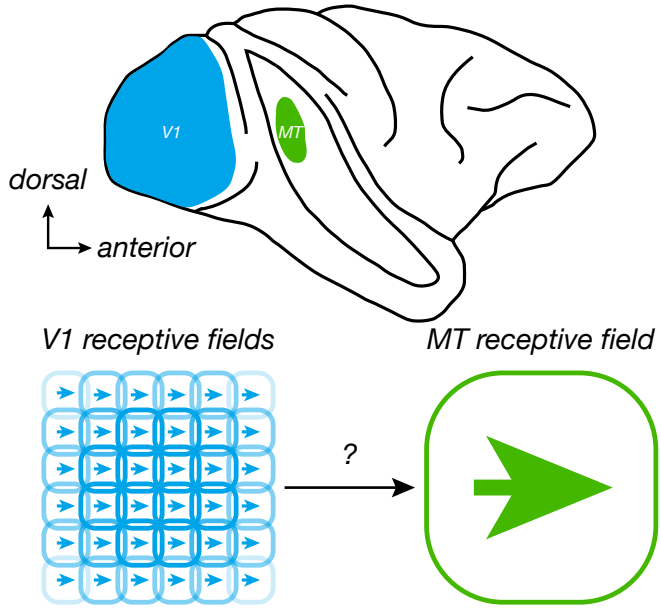

Figure 1.1: A lateral view of a macaque monkey brain with the temporal sulcus spread apart. Two visual areas that are central to motion processing are shown: V1 (in blue) and MT (in green). Many small directional V1 RFs converge to form a larger directional MT RF. I addressed whether the directionality of an MT neuron is limited to the scale of its inputs or whether it possesses an additional, more widely spaced, directionality. The arrows within the RFs represent the preferred directions for those cells, which are the same in all cases for illustration only.

Another difference is that MT RFs are roughly tenfold larger than V1 RFs at corresponding eccentricities [116]. This enlargement results from a tiling of visual space with convergent input from smaller upstream RFs. But MT RFs are not simply expanded V1 RFs. The preferred spatial frequency of an MT neuron is likely to include multiple cycles of an elongated grating, whereas a single cycle would be more likely preferred by a V1 neuron.

Motion processing in humans might be limited by the scale of directional tuning in MT, but humans are capable of seeing motion even when it occurs over wide separations of visual space, larger than the RFs of MT's inputs. This begs the

question whether directional responses in MT are limited by the scale of the RFs of the inputs to it or whether it forms an additional representation (Figure 1.1).

To test these possibilities, I isolated motion content that could be processed by V1 from that which might be processed in MT. I used a stimulus whose DoM was independent at two spatial scales. I defined 'local motion' to be that of a drifting grating covering a subregion of an MT RF. I defined 'global motion' to be that of a sequence of small patches of local motion, which were flashed in sequence from one side of the RF to the other at many different speeds.

1.2 Methods

1.2.1 Electrophysiology

I recorded from isolated single units in MT of six anesthetized paralyzed adult Cynomolgus Monkeys (*Macaca fasicularis*) using previously described methods [117]. All experiments were performed in compliance with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and within the guidelines of the New York University Animal Welfare Committee. Anesthesia was maintained throughout the experiments with a continuous intravenous administration of 4-30 μg/kg/hr of Sufentanil in dextrose-saline (4-10 mL/kg/hr). Vecuronium bromide (Norcuron) was also infused at 0.15 mg/kg/hr to prevent involuntary drifting of the eyes.

Gas-permeable contact lenses were used to protect the monkey's corneas. Supplementary lenses that were chosen by direct ophthalmoscopy were used to make the retinas conjugate with a screen on which all stimuli were presented. Refractive correction was checked by adjusting the lens power to maximize the resolution of

recorded cells. During experiments, each monkey was artificially respirated and body temperature was maintained with a heating pad. Vital signs (heart rate, lung pressure, EEG, ECG, body temperature, and end-tidal CO_2) were monitored continuously.

I passed quartz-glass microelectrodes (Thomas Recordings, Giessen, Germany) through a small durotomy within a craniotomy centered 15 mm lateral to the midline and 4 mm posterior to the lunate sulcus. The electrode was advanced at an angle of 20 deg from horizontal in a ventroanterior direction in the parasagittal plane. RFs were centered between 3 and 28 deg from the fovea, although the majority were between 4 and 12 deg. Signals were amplified, band-pass filtered (300 Hz to 10 kHz), and fed into a time-amplitude window discriminator (Bak Electronics, Mount Airy, MD). Spike arrival times and synchronization pulses were recorded with a resolution of 0.25 ms.

After an experiment was completed, the monkey was euthanized with an overdose of sodium pentobarbital (60 mg/kg) and perfused with 4% paraformaldehyde. I used standard methods for histological confirmation of the recording sites (Kohn and Movshon 2003). Identification of the recording locations was made through histological identification of electrolytic lesions that were made at suitable locations along the electrode tracks during the experiments. For that, I passed 1-2 μA of current for 2-5 s through the tip of the electrode.

1.2.2 Stimuli

Stimuli were presented at a resolution of 1024x731 on a gamma-corrected Eizo T550 monitor with a refresh rate of 100 Hz. The monitor was usually placed 80 cm from the monkey's eye, where it subtended about 22 deg of visual angle. The

mean luminance of the monitor was 33 cd/m^2. I used a 10-bit Silicon Graphics board to generate stimuli. They were presented to each neuron's preferred eye. The other eye was covered for the duration of the recording done with a given neuron. Stimuli were centered as close as possible to the center of a neuron's RF.

Local-global motion stimuli consisted of multiple local-motion pulses presented in sequence (Figure 1.2a). Each pulse contained a small, brief, spatially and temporally band-limited drifting grating. I used luminance-modulated raised-cosines in the x and y direction for spatial windowing and a raised-cosine for the temporal windowing. The long-axis of the local pulses was set to be orthogonal to the preferred axis of direction selectivity for the neuron. Local and global motion was presented in the preferred or null directions for the cell, giving four possible combinations of directions (Figure 1.2b).

I chose the width and duration of a single local pulse by measuring the response of a cell to stimuli where these parameters were parametrically varied. I chose values for these at which responses were strongly directional. Normal values were from 0.5 to 2 deg for width and 70 to 90 ms for duration. The spatial and temporal offsets between local pulses were chosen such that the range of global speeds were centered on the preferred speed of the neuron. Preferred speed was estimated in preliminary experiments.

I did not measure the RF widths of the inputs to the cells that I recorded from, which would have required significantly more work, and of a different sort. In separate experiments, I would need to determine which cells are projecting to an MT cell that I explored, perhaps with the low yield approach of determining which V1 neurons are antidromically activated by electrical stimulation of MT.

It is possible to say something about this from the literature, though, and a

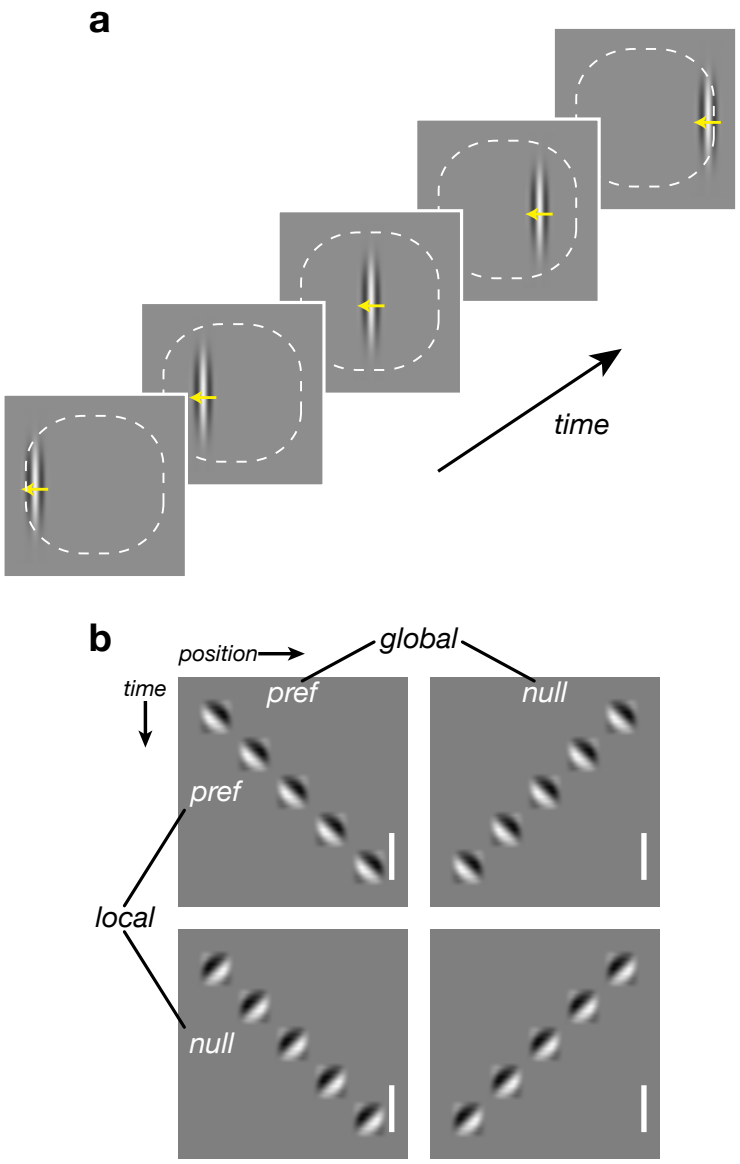

Figure 1.2: (a) A diagram of a sequence of frames of an apparent motion stimulus with local motion to the left (yellow arrows) and global motion to the right. The MT neuron whose RF is represented by the dashed white lines could prefer rightward motion, such that the motions of the stimulus are in the null direction locally and in the preferred direction globally. (b) Space-time plots of the local-global motion stimulus shows the four pairwise combinations of preferred and null local and global motion. Notice that the orientations stimulus content within each rectangular blip is either parallel to or orthogonal to the orientation of the overall sequence. The white bar corresponds to 100 ms.

simple, and often cited, relation is that MT RFs are tenfold larger than their V1 inputs [41, 118, 38, 119]. This is largely based on linear fits of the distribution of sizes across eccentricity in the two areas. In other words, it is not clear which of the cells with the smaller RFs in V1 are the inputs to the larger MT RFs at a given eccentricity. Also, the existing data has yet to incorporate variation in the relationship between the RF sizes of of the two areas across different cortical layers. Nor is it clear how estimates of the relationship varies with the stimulus used, or with regard to the extent of overlap in the input RFs. There is also asymmetry in the length of the RF in different directions of visual space, which relates to the tuning properties of the cell, which has not been considered.

That said, the width of the local pulse that I used was chosen in relation to the spatial frequency tuning of a cell and to ensure that five of them, when presented adjacent to one another, would fit within the cell's RF. The width of the local pulse was, therefore, likely to be on the order of one or two RF widths of the V1 neurons projecting to MT, but more work would be needed to firmly establish that.

1.2.3 Data collection

I determined the optimal direction, spatial and temporal frequency, and size of a high contrast drifting sine-wave grating for each neuron. I used optimal parameter values for all subsequent local-global experiments. I included all responsive neurons although I excluded a few that were unresponsive to spatial frequencies less than approximately 2.8 c/deg. This ensured that I would be able to present at least five local pulses within a cell's RF since their spatial offset was equal to the optimal spatial wavelength and I always presented at least one full cycle.

Next I varied the width and duration of a single local pulse in separate exper-

iments. For these and all subsequent experiments, I collected five trials for every condition, with five repeated presentations of the stimulus for each trial. I chose parameters for the local pulses such that they were as narrow, long and brief as possible.

In the main set of experiments, I varied the temporal offset between local pulses. I presented stimuli with global speeds that were much faster and slower than a cell's preferred local speed. Within the series of eight global speeds were three steps of doubling the speed and three steps of halving the speed from middle point matched to the local speed. I included an internal control in which one of the global speeds was zero, so that all blips were presented simultaneously. The spatial separation for each blip in this series was equal to the blip width, so that there was no spatial overlap between individual blips.

I followed up with two controls. In one, the temporal frequency of the local pulses was set to zero, nullifying the local motion. With that difference, I ran the same series of eight temporal offsets as before. In the second control, I replaced the local motion pulses with a narrow black bar. In a matrix, I varied the spatial and temporal offsets of these drifting bars. The temporal offsets were of 10, 20, 40 and 80 ms. The spatial offsets ranged from approximately 1 to 2.5 deg in four roughly equivalent intervals. For the two smaller spatial offsets, I presented seven pulses instead of the normal five.

1.2.4 Data analysis

I computed the mean instantaneous firing rate in 1 ms bins across passes within a trial and across trials for all cells. I plotted peristimulus time histograms (PSTHs) for all cells after smoothing them with a Gaussian filter with a

3 ms standard deviation. In addition, I computed response rates for all cells as the mean firing rate across passes and trials within brief temporal intervals. The intervals were shifted relative to presentation of the stimulus. I used an onset latency of 25 ms and an offset latency of 45 ms. When the global speed was slow enough that the intervals did not overlap in time, I excluded the periods outside of them. When they overlapped, I considered the window from the start of the stimulus plus the onset latency to the end of the stimulus plus the offset latency, a single period. I subtracted the spontaneous rate, which was drawn from a 600 ms blank period and averaged across all possible presentations.

1.2.5 Metrics

To quantify the extent of directionality at the two spatial scales, I initially computed a simple directional index as

$$1 - \frac{N}{P}$$

with summed response rates. For the local directional index, for example, I summed response rates for the two directional conditions that contained null local motion. I did the same for the conditions with preferred local motion. I computed the corresponding global directional index using a similar ratio of response rates. These were useful for getting a sense of the scale at which cells were directional, but I sought a more sophisticated metric that captures the relative dominance of directionality at one scale over the other.

I computed two metrics, one for each scale, and took the difference between them. 'Locality', one of the two, was the difference of the preferred local and null

responses, where each of these is the average over the two global conditions, divided by the sum of the absolute values of the same responses, the preferred and null local responses. 'Globality' was computed in the same way, as the difference of two response rates divided by the sum of their absolute values. I termed the difference between locality and globality 'local dominance', which captures the strength of directionality to the local motion over the directionality to the global motion.

Local dominance roughly corresponds to the directed distance from the equality line for points in a space of local directional index versus global directional index. But local dominance differs from that distance in that it is bounded from -2 to 2 since locality and globality are each bounded from -1 to 1. The absolute values in their divisors negate the influence of suppression below baseline. For example, response rates of 100 and -5 imp/s for preferred global and null local conditions have a locality of 1. Response rates of 100 and -20 imp/s have the same locality. The local directional index for these rates would be 1.05 and 1.20 respectively. Local directional index, like global directional index, will scale with greater suppression below baseline. Local dominance values greater than 1 can only result, therefore, when there is significant directionality to the local motion and responses to the global motion that are greater for null conditions than for preferred ones.

1.2.6 Model

I constructed a simple motion energy model to explore whether cells' responses were determined by the frequency content within the local-global stimuli. Predicted responses were computed by convolving a cell's spatiotemporal RF with a local-global stimulus. To get a cell's sensitivity profile in the frequency domain, I assumed separability and took the product of the cell's spatial and temporal

frequency tuning. There is some evidence that MT cells' tuning is not separable, that they are partly tuned to velocity. But I considered this approximation a reasonable starting point, which should be able to predict responses to my stimuli. I computed the same directional metrics for the simulated response rates as for the physiological ones.

1.3 Results

1.3.1 Scale of directionality

I recorded from 50 neurons in MT of six anesthetized paralyzed macaque monkeys. I began my experiments with a single cell by defining the properties of the local motion. I ran two experiments to parametrize a single centered patch of local motion such that a neuron was directionally selective to it. From the following experiment in which six different durations were presented, I chose the shortest, or just longer than the shortest, duration that would elicit a directional response (Figure 1.3a). Most neurons were directional for durations longer than 30 ms. And although their responses to preferred local motion increased only slightly for longer durations, their responses for null motion were increasingly suppressed up to approximately 70 ms.

I reasoned that if there is selectivity for global motion it would be most easily detected when there is a significant response to the local motion. I also assumed that global motion would be more likely to have an effect when its speed is the same as a neuron's preferred local speed, although the actual relationship between preferred local and global speeds might be quite different. I could not be sure either way and I explored a broad range of global speeds.

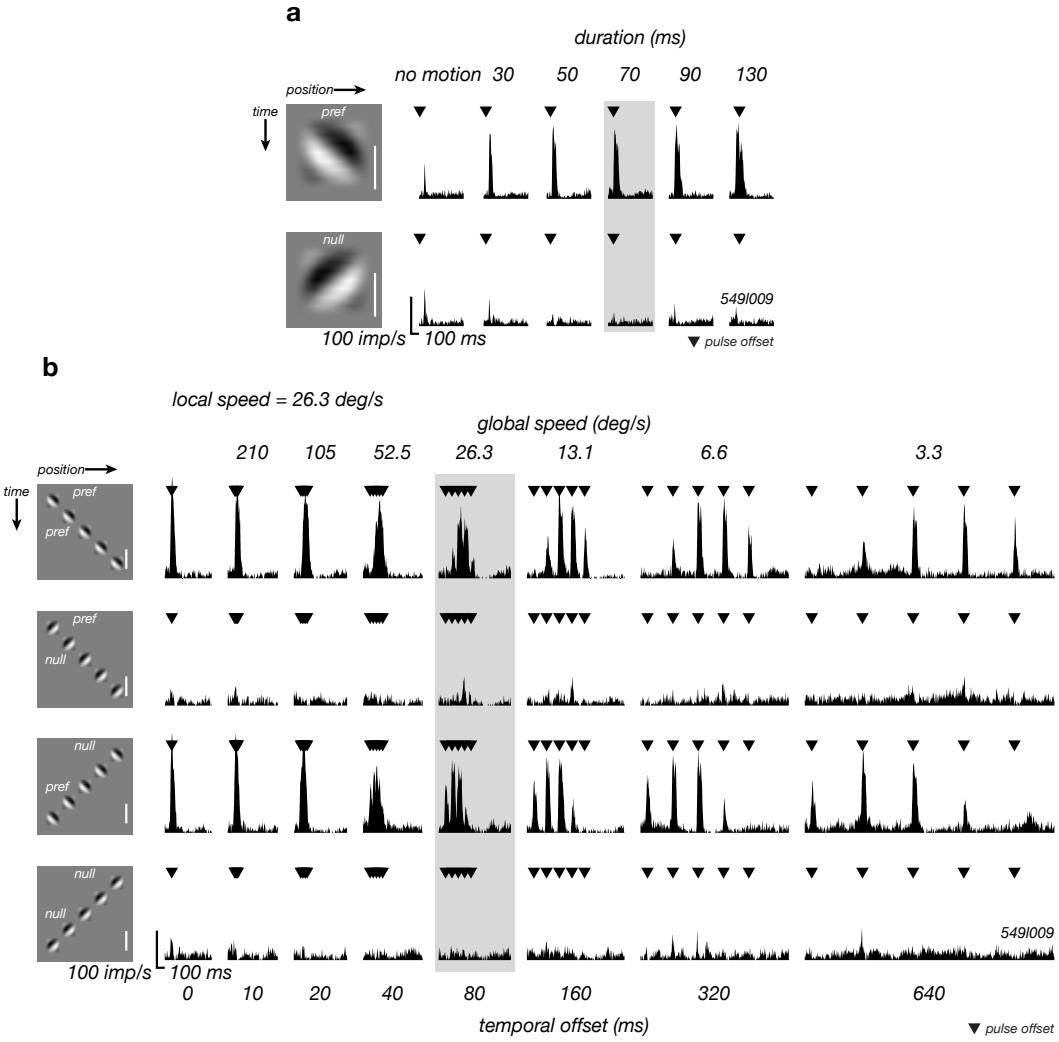

Figure 1.3: PSTHs of the responses of an MT neuron to a single local pulse with six different durations and to a full set of local-global motions with eight different global speeds. The space-time plots on the left are for the duration or global speed condition highlighted in gray. The white bar in the space-time plots is for 100 ms. (a) The top row is for motion in the preferred direction and the bottom in the null direction. For all of the main local-global experiments, I set the duration from this series, by choosing one for which there was a clear directional response. (b) The four rows are for different directions of the stimulus. Rows 1 and 2 are for global preferred. Rows 1 and 3 are for local preferred. Global speed decreases from left to right. The cell responded strongly to conditions with preferred local motion at all global speeds.

The responses of a typical MT neuron to eight global speeds and four directional conditions of local-global motion is shown in Figure 1.3b. Each column is a set of PSTHs for the four directional conditions for a single global speed. Global speed increases from right to left. For the slowest global speed (far right column), the neuron responded when the local motion was in the preferred direction (first and third rows), regardless of the direction of the global motion and was equally suppressed otherwise (second and forth rows). The lack of influence of the global motion for this condition might not be very surprising since the global speed was very slow relative to the preferred local speed. The temporal offset in this case was 640 ms such that each blip was presented as essentially isolated in time, progressing slowly from one side of the RF to the other.

In the condition where the global speed matched the neuron's preferred speed (26.3 deg/s), this neuron responded only to the local motion and was unaffected by the global motion (Figure 1.3b, gray highlight). This neuron was like some others in that its response to the first and second blips of local motion for one global motion direction was similar to the forth and fifth blips for the condition for which the global motion was in the opposite direction. This might result from the stimulus being slightly off center with respect to the RF. There was also some variability in latency across different directional conditions in some other neurons. Whatever the explanations are for these subtleties they are not evidence of a representation of global motion.

At the fastest global speeds (52.5-210 deg/s), the blips overlapped in time and there was a continuous response to the preferred local motion. The responses at those speeds were similar to the ones at slower global speeds: the enhancement and suppression to preferred and null local motion was unaffected by the global motion.

Also, the responses increased with global speed, which probably results from there being a greater number of blips within the RF at a given time. It seems as though there is incomplete summation of the responses. The control condition in which five blips were shown at once (far left column) supports this. These responses, like the others at fast global speeds, are less than what one would expect from summing the response to each blip in isolation, such as at the slowest global speed (far right column). In general, the responses of the neuron I have focused on were similar to those of the population.

Figure 1.4a summarizes the response rates for the seven non-zero global speeds for the example MT neuron. The response is much greater, and roughly equal, for the conditions with preferred local motion. The response to preferred motion increases as the global speed increases. I quantified the influence of global motion by first comparing the response rates for local preferred and null motion with preferred and null global motion. Then I computed a single metric based on the difference between the directionality to the local or the global motion. Local dominance ranges from -2 to 2 and is consistent with purely local directionality at 1 and purely global directionality at -1. A value of 1 does not, however, mean that the selectivity is only determined by the local motion since many different combinations of the two metrics on which it is based can result in the same local dominance. Figure 1.4b shows the local dominance values for these speeds for this neuron. They are near 1, consistent with it being primarily influenced by local motion.

Local dominance for all cells is nearly always greater than 0 and is centered near 0.7, meaning most cells are directional to the local motion (Figure 1.4c-e). There are a small number of cases that have local dominance values near, or slightly less

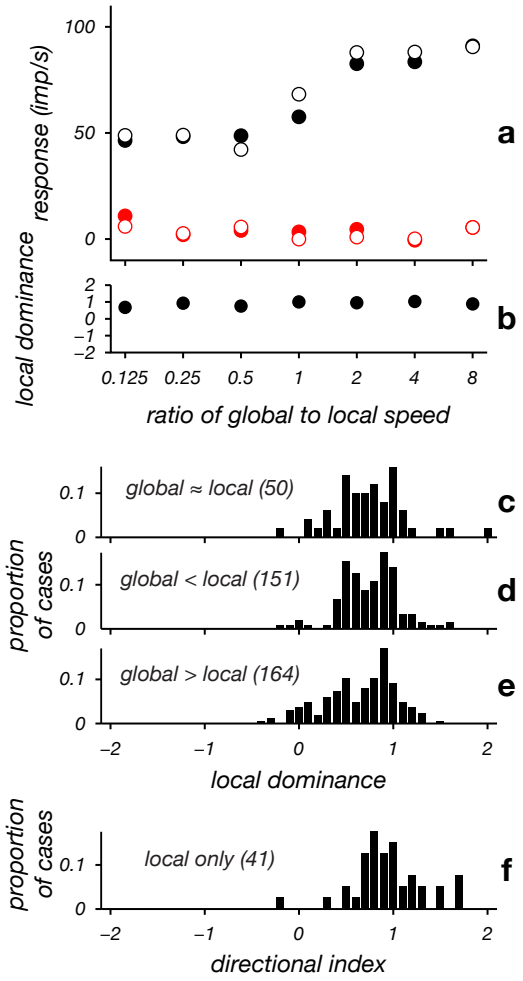

Figure 1.4: (a) Response rates for stimuli with local-global motion for the example MT neuron I show throughout this chapter. The points are: black for preferred local, red for null local; filled for preferred global, hollow for null global. Global speed increases from left to right. The cell was directional for local motion. (b) I computed 'local dominance' from these response rates. It is more positive when a cell is relatively more directional for local motion and more negative when it is more directional for global motion. (c-e) Histograms of local dominance for three groups of global speed for a population of MT cells. (f) For comparison, I computed a simple directional index based on the ratio of the responses for null and preferred motion for the control condition without global motion. At all global speeds, nearly all cells were directional for local motion.

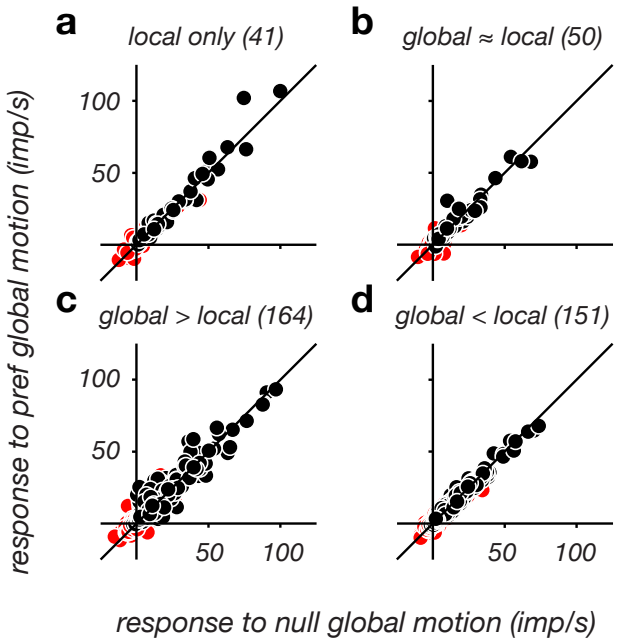

Figure 1.5: Response rates for a population of MT neurons grouped by global speed. The abscissa and ordinate are the response rates to null global motion and preferred global motion respectively. The points are: red for null local, black for preferred local. The responses to null and preferred local motion were unaffected by the direction of the global motion for all global speeds.

than, zero. I reviewed the PSTHs of this subset of neurons and found them to be essentially unresponsive to all of the conditions for global speed and motion direction. This suggests that their local dominance values are primarily the result of noise. Figure 1.4f is for the case where there was no global motion, so I show the local directional index (LDI) instead of the local dominance, since the latter does not make sense in this case. LDI, in contrast to local dominance, is unbounded.

Figure 1.5 shows scatter plots of the response rates for null versus preferred global motion. Local preferred response rates are shown as black points and local null response rates shown as red points. The four plots are for different groups of global speeds across all neurons or all conditions for all neurons. Figure 1.5ab

is for a single global speed for each neuron, whereas Figure 1.5cd includes several cases that are within those speed ranges. It is clear that for all neurons, and at all speeds, neither the preferred nor the null local responses are influenced by the direction of the global motion.

1.3.2 Responses to global motion only

The absence of a response to global motion in MT could be explained by inability to significantly drive a neuron and in a sense my results seem to support this explanation. Another possibility is that the response to local motion dominates the global one, that it masks an underlying response to it. To check for this, I measured responses to global apparent motion without local motion. Figure 1.6a shows PSTHs of the response to global motion in the example MT neuron. As for Figure 1.3, each column shows PSTHs for a single global speed, which decreases from left to right. The direction of the global motion is indicated by the orientation of the content in the space-time plots.

The neuron responded equally to motion in either direction for all speeds. It was not directional for global motion. This suggests that there is not an underlying directionality for global motion that is masked by local motion. The responses for these conditions were less than the responses to preferred global motion with preferred local motion, and greater than the responses to preferred global motion with null local motion. This suggests that local motion enhances and suppresses what would otherwise be non-directional responses. Also, the responses decrease as the global speed decreases for the three global speeds that are faster than the condition matched to the preferred local speed, as they do for the two local preferred conditions for the main local-global series.

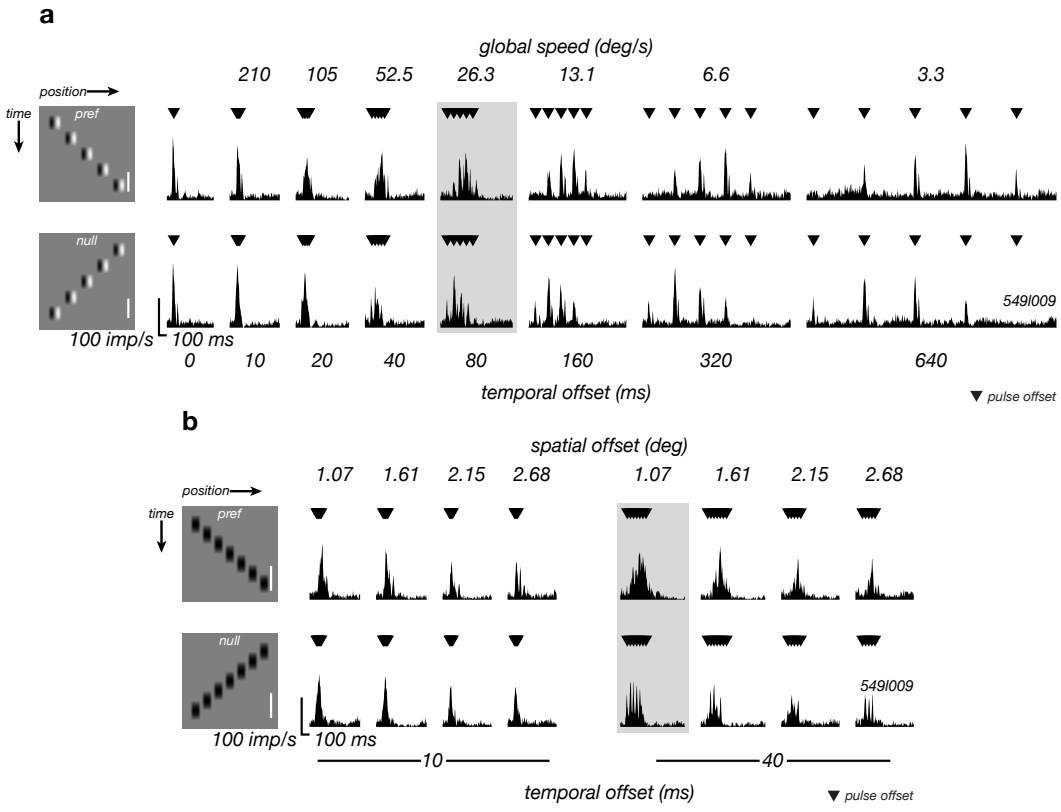

Figure 1.6: PSTHs of the responses of an MT neuron to for two control experiments in which the stimuli contained no local motion. The space-time plots on the left are for the conditions highlighted in gray. The white bar in the space-time plots is for 100 ms. In both sets of PSTHs, the top row is for motion in the preferred direction and the bottom in the null direction. (a) Global speed decreases from left to right. The responses were not directional. (b) I varied the spatial and temporal offsets between the apparent motion pulses. At these temporal offsets, the cell's responses were not directional at any spatial separations.

I did a second control series in which only a single scale of motion was presented. Low spatial frequency gratings (black bars) were presented at a broad range of speeds. I did this to check whether there is any directionality in the range of the spatial separations of the global motion I presented in the previous experiments. These stimuli were similar to the ones used by Mikami et al. [120, 121, 122]. Figure 1.6b shows PSTHs for the example MT neuron for different spatial offsets at two different temporal offsets. These were a subset of a larger grid of all combinations of four temporal and four spatial offsets that I presented. Since the spatial offset increases from left to right, the speed increases from left to right. The space-time plots on the left show the direction of the motion. For this neuron and this range of offsets, there was no clear directionality for any of these conditions. This further affirms that directionality in MT neurons is limited in the spatial extent over which it can occur.

MT neurons were generally not very directional for any of the spatial and temporal offsets that I presented. Figure 1.7a-e shows histograms of directional indexes for all neurons for these black bar stimuli and for the previous set which were like the local-global stimuli but lacked local motion. In a few cases, there was some weak directionality, which probably resulted from the stimuli being just below the upper limits of offsets to which directionality exists. The median global DI was 0.138 with the black bar stimuli, consistent with weak directionality. Any directionality tended to occur when the spatial offset was between 0.5 and 1.5 deg, with little to no directionality beyond this. These spatial offsets are less than the ones typical of my local-global stimuli.

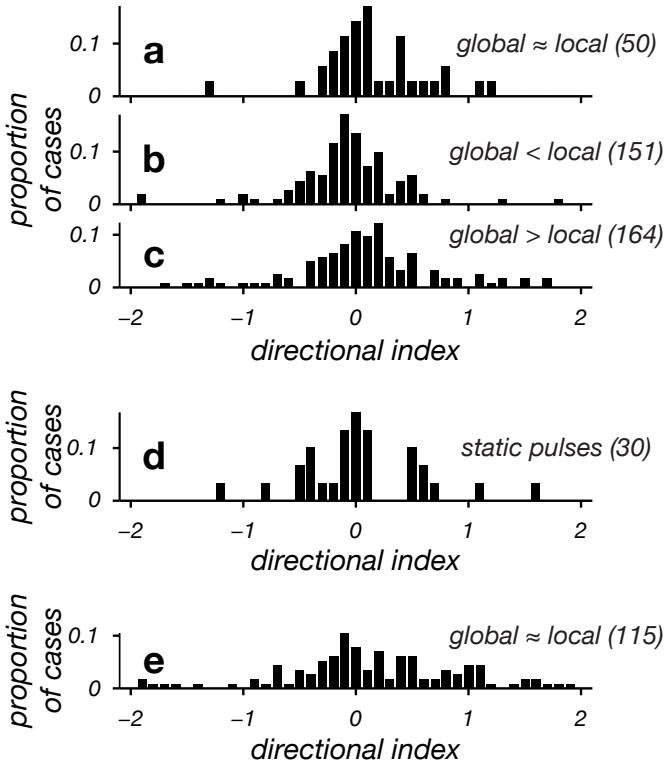

Figure 1.7: (a-c) Histograms of directional index for three groups of global speed for a population of MT cells. (d) A histogram of directional index for flashed local pulses, so no local or global motion. (e) A histogram for the conditions of the black bar stimuli in which the global speed was roughly equal to the local speed. At all global speeds, cells were not directional for stimuli that lacked local motion.

1.3.3 Motion energy model

The standard motion-energy model of MT neurons includes space-time filters that estimate motion-energy in a specific direction [59, 95, 62, 123]. Motion energy, therefore, drives the responses of MT neurons. I hypothesized that the pattern of directional responses that I observed could be explained in terms of the overlap of the motion energy of a stimulus with a cell's spatiotemporal response profile. I developed a simple model to test this. I do not attempt to simulate all of the characteristics of MT neurons.

Figure 1.8a shows space-time plots for the condition in which the global speed matched the preferred local speed. The direction of the global motion corresponds to the orientation of the group of five local pulses. The plots to the right of the space-time plots show the corresponding stimuli in the frequency domain. The red overlay shows the neuron's response profile to different spatial and temporal frequencies. The rightward preferred motion in the first row of plots in Figure 1.8a has circular blobs in the first and third quadrants with intensity striations within them. These energies fall mostly within the neuron's region of sensitivity. Global motion is a sampling of the local pulses and it corresponds to the striations of the local pulse energy. The orientation and spacing of these striations covaries with the speed of the global motion. They are more vertical and more widely spaced the faster the global speed.

The spectra for the case in which the stimuli are the same as the local-global, but without local motion, are centered on the abscissa (Figure 1.8b). The black bar stimuli should be more likely to result in directionality at the smallest temporal and spatial offsets, where there is the greatest separation between the components in the frequency domain (Figure 1.8c). In that case, orientations that correspond to

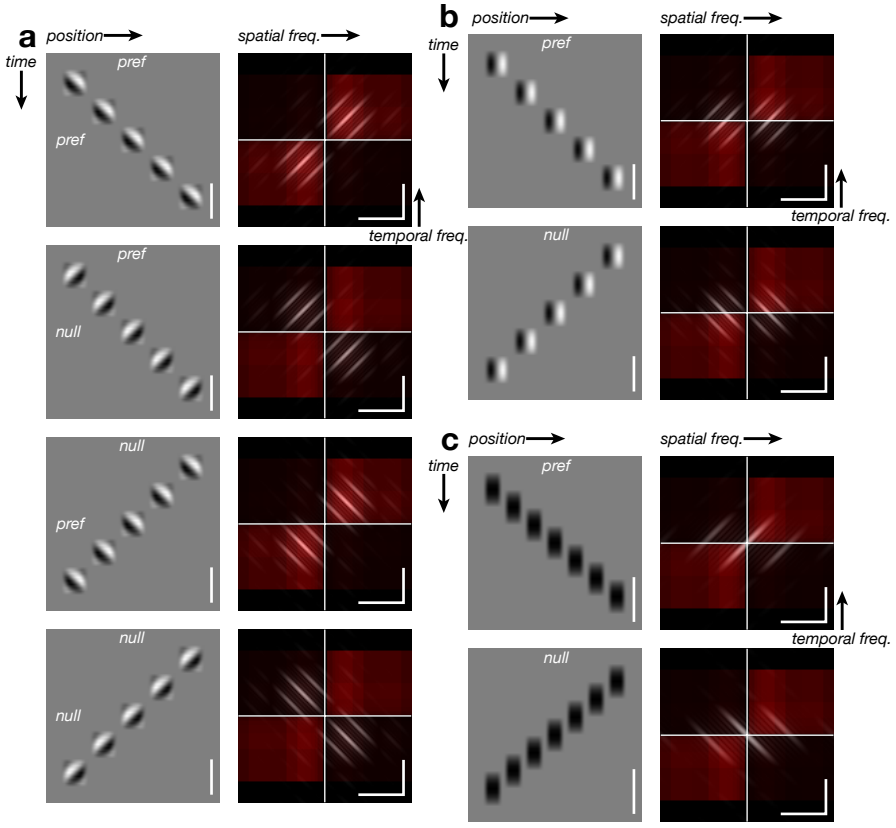

Figure 1.8: The gray colored plots are space-time plots of stimuli shown in previous figures. The white bar in the space-time plots is for 100 ms. The rightmost columns of plots are the Fourier transforms of the stimuli to the left of them. They have a red overlay that corresponds to the spatiotemporal frequency sensitivity of an MT neuron. For the frequency space plots, the vertical segment of the white bar is for 20 Hz and the horizontal is for 4 c/deg. (a) Plots for the condition of the series in which global speed was varied where global speed matched local speed. The local motion loosely corresponds to the circular energy regions in cross-quadrants. The global motion loosely corresponds to the striated sample lines within the circular regions. (b-c) When there is no local motion, the energy is centered on the abscissa. The direction of the global motion changes the orientation of the striations, although the amount of energy in the cell's RF remains about the same.

motions in different directions could result in different amounts of motion energy in the sensitivity window, and therefore differential responses. These controls make clear that the direction of the global motion corresponds to orientation, which may not influence the amount of energy within the sensitivity region (Figure 1.8bc). Therefore, within this model, the responses are unlikely to differ with global motion.

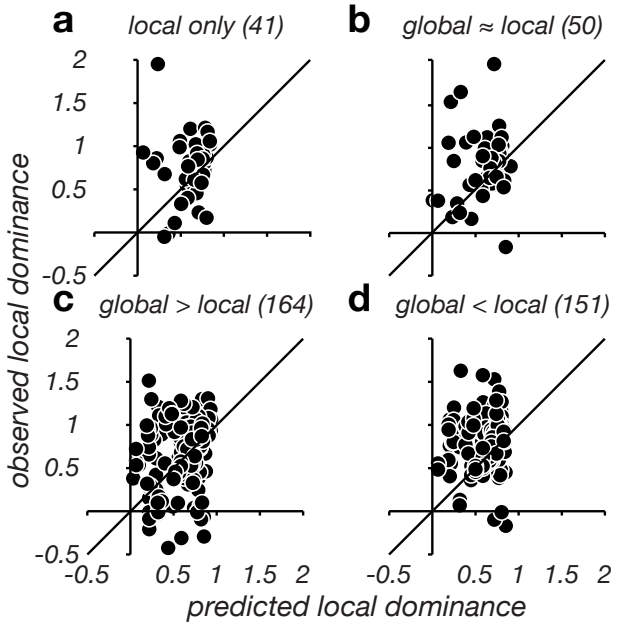

Figure 1.9: Local dominance for a population of MT neurons from a motion energy model and for actual data grouped by global speed. The model predicts the correct sign of local dominance, but not the range of variability seen in the real data.

I computed simulated responses for all MT neurons. Local dominance was computed in the same way as for the electrophysiological data. The model could not predict negative response rates, in part because there is not a clear way to simulate the baseline response. It also does not include suppression, such as from null motion. The range for the predicted local dominance is 0 to 1, since the model is selective for the local motion and does not give negative response rates, which can

extend the metric higher than 1. Another possibility for extending the range is for the globality to be negative from a greater response to conditions with null global motion than those with preferred. The scatter plots in Figure 1.5 suggest that this is not a general phenomenon. Figure 1.9 shows scatter plots for the simulated data, grouped by global speed. For the condition in which there is no global motion, the model generates reasonable predictions of slightly greater than 0.5. The data from the physiology for this condition has some variation (orthogonal to the axis of the predicted response). When the global speed matched the preferred local speed the model also did a reasonable job of predicting the data.

Figure 1.9cd shows cases where the global speed is greater than and less than the preferred local speed. The local dominance from the physiology data has significantly greater variability. This is especially true for the set with faster global speeds. It is unclear what underlies this. It is possible that there are not only greater responses, but more variability in the responses when there is more stimulus in the RF. This may be an issue of partial summation of the added stimulus, which I mentioned previously. The speed of the global motion, and the amount of stimulus in the RF at a given time, do not affect the model. In general, the predicted local dominance values are similar, but less than the actual ones. Simulating trials and velocity-tuning should improve the predictions.

An additional point to be drawn from the predicted local dominance values is that they are mostly distributed in the same quadrant as the local dominance values from the electrophysiological responses. The model predicts that responses are directional for the local motion, which is consistent with what is found in MT neurons.

1.4 Discussion

My results show that MT neurons are unaffected by global motion. This implies that MT does not integrate the representations of motion that make up its RF. Instead it simply aggregates the feed-forward inputs it receives from V1 and other areas. I was able to approximately predict a MT neuron's directionality by computing the convolution of my stimuli with each neuron's spatial-temporal response profile. A cell responded whenever there was energy within the region of frequencies to which it was sensitive. This suggests that a neuron that is sensitive only to the global motion would respond differentially to the striated peaks within the circular patterns of energy for the local motion. These would be like large oriented filters in space-time, which, as a result of their scale, would be insensitive to smaller orientation changes within their slanted filter regions.

It would be interesting to know more about how the responses of MT neurons are affected by global motion when the strength of the local motion signal is reduced. I did not explore combinations of global speed with local speed other than the optimal one, which I estimated by finding the preferred spatial and temporal frequency in separate experiments with different stimuli. As I explained previously, I set the local speed to that which would elicit a maximal response. Choosing a different, non-optimal, local speed would reduce responses, although, doing so should not induce a response to the global motion.

The mean local dominance that I observed was 0.7, which could result from a partial influence of the global motion. I believe this to be unlikely, although it cannot be ruled out from my present set of results. On this point, Figure 1.4f shows a histogram of a simple directional index from the responses to the same local-global stimulus, but without global motion for the subset of the neurons that

I recorded from. Without global motion, all of the local pulses were presented simultaneously. The mean of the directional index for this condition is less than 1 and is widely distributed, similar to the data for stimuli with global motion. There do not seem to be any obvious differences between these distributions.

Also, nonzero responses to null motion will decrease local dominance. In other words, incomplete suppression from null local motion could shift the distribution of local dominance towards zero, regardless of the direction of the global motion. The red dots in the scatter plots in Figure 1.5 are for null local motion. They are not centered around zero, but extend out to slightly postive responses, even for the condition without global motion. These scatter plots suggest that there is not a general systematic influence from the direction of the global motion.

I believe that my results are not significantly influenced by the use of anesthesia for my recordings. Previous studies with local-global motion stimuli led to similar results with awake monkeys [46]. I have thouroughly analyzed the results of these experiments, which were done with two awake monkeys [46]. I have not presented those results here, although they are consistent with these results in that MT cells are not influenced by the direction of the global motion.

It is unlikely also that changes in attention, such as attending to the global motion, or completing a task that depends on it, would give rise to a selectivity for the global motion. Again, additional work would be needed to rule that out. But there seems to be no signal of global motion to amplify in order for there to be an effect of attention. Training monkeys to report the direction of global motion in local-global motion stimuli might be a first step in understanding this further. This monkey psychophysical data would be useful also to determine whether other brain areas form a representation to global motion and how those representations

correlate with what is perceived on different trials.

There have been few published reports that I am aware of that describe physiological responses to simultaneously presented visual motion in opposite directions at different scales. Most earlier work made use of stimuli with a single scale of motion, and looked at the spatial and temporal offsets at which cells were directional [120, 121, 122, 42]. That said, I would like to summarize a few observations from previous studies that motivated my experiments, and which, in particular, suggested that MT might represent global motion.

Humans (and monkeys) have at least two different motion-sensing mechanisms: one that is capable of representing motion at spatial and temporal scales that are larger than those of the other [124, 125, 44, 45]. Early support for this came from apparent motion in foveal random-dot kinematograms. The standard cutoff for the 'short-range' process was 0.25 deg and with interstimulus intervals of less than 100 ms [124]. Subsequent studies showed that this depended on eccentricity as well as other factors, such as the size of the stimulus [125]. Regardless, the long-range process from which humans can perceive motion can be separated by as much as 10 deg and as much as 300 ms. The ability to perceive motions at these scales suggests that there are neurons whose responses are directly related to such motions.

MT neurons, in part because of their large RFs, seemed like a reasonable candidate for representing the long-range motion. Mikami and colleagues suggested that MT could represent motion at spatial separations that are significantly larger than those in V1 [122, 120, 121]. To be clear, these separations, although larger, were nonetheless still associated with the short-range process. They suggested that, among other things, the expansion in MT could facilitate directional selectivity at faster speeds. The expansion could be formed in part by indirect projections, pos-

sibly through V2 or V3, where the RFs are larger than those in V1. More recent work by Churchland et al., however, suggested that there are not major differences in the scale of directionality in V1 and MT [42]. This led them to conclude, as I have, that directionality in MT is inherited from V1. Because the global motion in my stimuli was defined by sampling the local motion, the motion energy of it had a specific shape that could be distinguished from the local motion. This is not the case for measurements of the maximum spatial and temporal offsets with random dot stimuli.

Some unpublished psychophsyical results suggest that the presence of global motion biases judgements of the direction of local motion. This suggests that signals for both motions converge. There are also psychophysical results that show that they can be cross adapted, which also suggests a shared representation. One way for cells to represent global motion would be to form a motion detector from inputs in which the directionality for local motion has been averaged out. This would seem most likely to occur in a step from smaller RFs, such as ones capable of representing the position of an individual blip in my stimuli, instead of larger RFs, like those in MT. This seems unlikely to occur in one of the primary downstream targets of MT [126]. Neurons in the ventral stream that are able to track objects moving over large distances may represent global motion. Either these areas or their targets could be points at which convergence with local motion could occur.

A study by Ilg and Churan is relevant to mine [126]. They trained rhesus monkeys to do a direction discrimination task for three different types of motion: Fourier motion, drift-balanced motion, and theta motion. In general, monkeys were significantly above chance level in discriminating all of these types of motion. This suggests that monkeys have mechanisms for representing second-order motion and

that they can perceive it, both for drift-balanced motion and for theta motion. This further supports the possibility that monkeys are capable of seeing global motion, as I have defined it, which is a form of second-order motion.

Additional experiments could be done to establish that monkeys can see long-range apparent motion, or what I have called global motion. There is little evidence that suggests they cannot see this motion, although it has not been thoroughly explored. The global motion that I presented has spatial limits that are significantly less than the upper limits for perceving long range motion, which operates over many degrees [127, 44]. The temporal offset may also be up to 500 ms, which is longer than nearly all of the temporal offsets that I explored [127, 44].

The perception of the direction of local and global motion seems to be influenced by the eccentricity at which they are presented. It is not known why this is? My data do not suggest that this variation with eccentricity is reflected in the tuning properties of MT neurons. Local dominance, for example, is roughly constant across cells whose RFs are at different distances from the fovea.

Ilg and Churan recorded responses from MT and MST during discrimination of the different types of motion [126]. They found that neurons coded only for stimulus movement if the motion stimulus was separated from the background by luminance or flicker, Fourier and drift-balanced motion. If these cues were absent, which was the case for theta motion, the neurons did not code for direction of motion. They concluded that MT and MST are not the final cortical stages that are responsible for forming percepts of motion.

I explored a different question, about the scale at which MT represents motion, and whether it computes a directionality beyond what is formed in V1. But my results are consistent with theirs in that, if one assumes the monkey can per-

ceive global motion, then the responses of MT neurons are not consistent with the percept. The neurons that represent theta motion, for example, might also represent long-range apparent motion. In both cases, they might be insensitive to the fine-scale motion within the stimulus.

Majaj and colleagues also tested whether MT neurons integration motion signals across space witin their RFs [128]. They presented drifting gratings in two separate patches, and compared the responses of cells to conditions in which the drifting gratings were superimposed within the same patch. Cells lost their selectivity to pattern motion when the gratings were separated in visual space. This is consistent with my results, in that it suggests that MT is not integrating across space within its RF. But my experiments tested not only whether signals are combined across space, but whether MT neurons are sensetive to the temporal order and speed at which activation in space occurs.

Chapter 2

Speed-dependent plaid perception, the interplay of two preferences

A drifting plaid can be perceived as a coherently moving pattern or as two overlapping transparent patterns sliding over one another. Previous studies have identified different parametric manipulations that can affect which percept is more likely to be seen, but they have not explained why. To learn more about the way the system decides, I measured the perception of drifting square-wave plaids with a forced-choice task. I presented plaids with 100 different combinations of component and pattern speed. For a range of intermediate component speeds, transparency was more likely for plaids whose pattern speeds were two to three times their component speeds. In this range, the percepts seem to follow the "angle rule," which is to say that transparency dominates when the pattern speeds are faster than some multiple of the component speeds. Plaids with faster or

slower component speeds than the range at which they were sometimes transparent, generally appeared coherent through the fastest pattern speeds. I hypothesized that the influence of the speeds on the percept arises largely from a competition between preferences to see slow and singular motion (i.e., coherent). I formulated a Bayesian model that imparts these preferences as prior distributions. The model has separate likelihood functions for each component. Simulated percepts from the model were well-matched to my experimental data. Changes in either prior influenced the pattern of percepts across speeds and may underlie inter-subject variability.

2.1 Introduction

A system that represents visual motion encounters an ambiguity at the level of local motions known as the 'aperture' problem. It applies not only to the visual system, which can represent only a finite region of visual space, but also to the receptors within the system, whose RFs encode a small region of the space [2, 129]. Consider the motion of a grating drifting within a circular aperture. Its motion is inherently ambiguous since any components of motion parallel to its orientation produce no observable change in the image intensity pattern [7, 130, 3, 8, 4] (Figure 2.1a). This is illustrated in the plot of the three relevant velocities of a square-wave plaid in 2D velocity space in Figure 2.1b. For each component, velocities along the dashed 'constraint lines' are possible given the visual input for the motions of the square-wave gratings shown in Figure 2.1a. The green vectors emanating from the origin are the shortest (slowest) possible that connect to the constraint lines. These normal velocities are consistent with what is perceived for those gratings

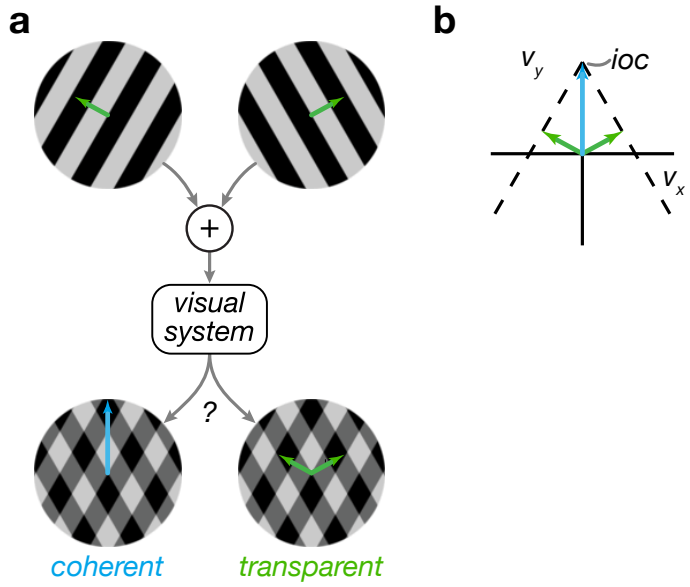

Figure 2.1: The perception of a drifting square-wave plaid and its velocity space representations. The colored vectors represent motion in a direction given by the vector's angle at a speed given by the vector's length. (a) An upward drifting plaid is represented by an observer's visual system. The plaid's motion is ambiguous: it can be perceived as a singular rigidly moving pattern (coherent percept in blue) or as transparent gratings sliding over one another (transparent percept in green). (b) Representations of the plaid's motions in 2D velocity space. The dashed lines indicate the possible velocities consistent with the windowed component and pattern motions of the stimulus. IOC is a close approximation of what is seen in the coherent case.

when viewed by themselves.

The aperture problem is resolvable by a global process when additional information is available. For the square-wave plaid shown in Figure 2.1a, the perceived velocity, when it is perceived to move coherently, is the one that is consistent with the intersection of the constraint lines of the two components. This is the IOC solution (Figure 2.1b). It is the unique 2D translational motion that is consistent with the motion of both gratings. Results from many physiology and computational studies suggest that the human and non-human primate visual systems

implement a similar integrative process for representing global motion. The existence of pattern direction-selective cells in MT supports this. The perception of plaids with certain stimulus configurations may be inconsistent with the IOC solution [131, 10, 132, 102]. For example, when the pattern direction is outside of the narrowest sector formed by the directions of the component motions, the perceived DoM may be biased towards the direction of the component motions. Also, the perceived pattern speed is often slower than what is predicted by IOC [133]. IOC is nonetheless a reasonable approximation of the pattern velocity for many configurations of plaids, including those that I explored.

I have focused on integrated motion thus far, but the humans can form representations of multiple overlapping objects in complicated real world scenes, which can move in different directions and at different speeds. In computing the global motions of overlapping objects, the visual system must not only integrate local motion signals. It must also segment them when it infers that the motions are for different objects. It must not simply 'solve' the aperture problem whenever it has sufficient constraints, but must decide what to integrate and what to segment on the basis of the input it receives. Braddick offered an example of this, which is that of a moving animal whose markings may provide it camouflage against the vegetation of its habitat [134]. Normally any markings on the animal do not move with respect to the background and, therefore, the animal is not detected. But the effect of the camouflage would be lost were the markings to move independently with respect to the background, in which case the animal would be detected as a separate object. Independent motions in overlapping portions of the scene can lead to the differentiation of one object from another, or from one object to other portions of the scene.

Plaid stimuli have been used to study the conditions for which segmentation or integration occurs. For plaids with different stimulus configurations, such as different sizes, spatial or temporal windowing, these manipulations have been shown to decrease coherence: increasing the speed of the components [7]; increasing the angle between the directions of the component motions [7, 14]; increasing the difference between the spatial frequencies, contrasts, or hues of the components [7, 135, 136, 137, 13]; adjusting the luminance cues at the component intersections so that they appear to overlap in depth [18, 15]; and increasing the duration of the stimulus presentation [138, 136].

Known relationships between a plaid's stimulus configuration and how it is perceived to move are mostly based on measurements of which percept occurs in forced-choice tasks for a few combinations of its possible parameter values. This leaves open the possibility that within other regions of the parameter space the relationship between a specific parameter and the likelihood of either percept could be insignificant, or reversed. For example, increasing the speeds of a plaid's components has been shown to increase transparency, but is this true at all speeds, or does it vary depending on other parameters, such as angle, spatial frequency, contrast, etc.? Is there still an effect if the angle between the directions of component motions is very small, so that the pattern speed is only slightly faster than the component speeds? Limiting the parameter space additionally makes it more difficult to unify the results of different studies into a coherent framework. I believe these uncertainties should be addressed not only by expanding the parametric region within which measurements are made, but by linking the psychophysical observations to models of how motion percepts are formed.

That said, despite nearly 30 years of investigations of drifting plaids it is

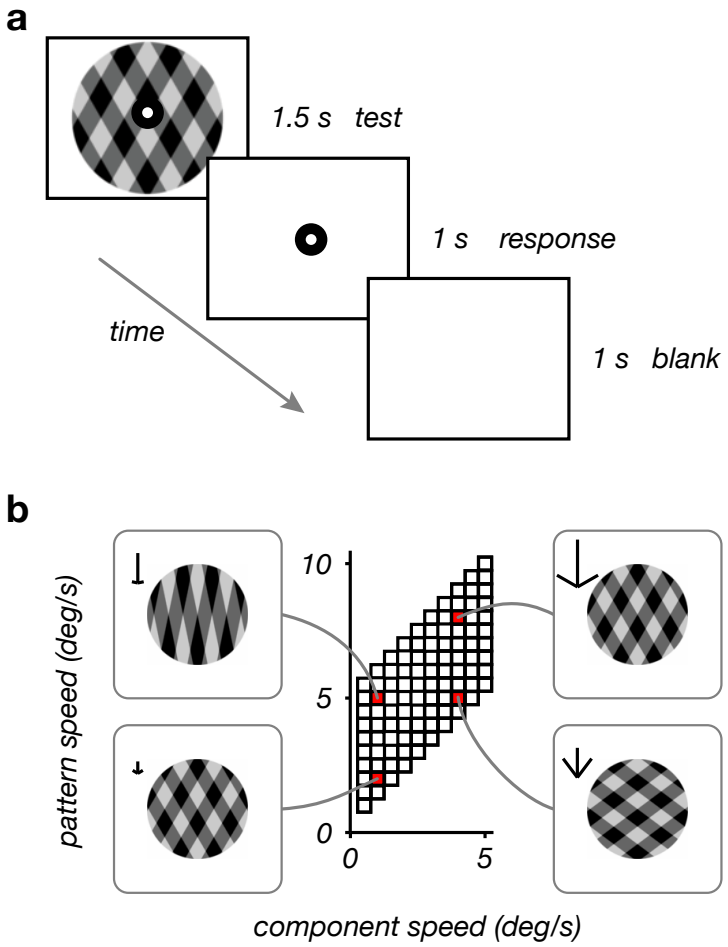

Figure 2.2: 2AFC protocol for measuring plaid perception and sampling in component and pattern speed space. (a) An upwardly moving square-wave plaid was presented for 1.5 s. This was followed by a 1 s response period during which subjects indicated whether the plaid appeared coherent or transparent. (b) Plaids were presented with 100 different combinations of component and pattern speeds. Each combination corresponds to a square in this space. Plaids represented in the lower left corner of this space, for example, have slow component and pattern speeds. This is indicated by the small vectors in the inset of the plaid in the lower left, for which length corresponds to speed and angle to direction. Four combinations of speeds (the red squares) are highlighted by showing a representative frame of the stimulus for the corresponding speeds.

presently impossible to state the conditions under which either percept will occur. Moreover, we cannot account for why a given manipulation effects the probability of either percept. There is no general model for how the visual system forms either percept. I conducted a set of psychophysical experiments to examine the joint effects of component and pattern speed of drifting square-wave plaids on how they appear to move. I measured how plaids appeared to move at many regularly spaced combinations of these speeds. The experimental design that I used is illustrated in Figure 2.2a. It was clear from my data that both speeds influenced the percept, so that coherence or transparency was dominant within different regions of the space. I hypothesized that the relationship between a plaid's speeds and the dominance of either percept is the result of the interplay between visual preferences to see a single pattern of motion and to see slower speeds. These preferences arise naturally in the context of a Bayesian decision model, which I developed to account for these results.

Bayesian models of visual motion begin with the idea that our perceptions can be explained as a combination of probabilistic inferences about motions in the world with prior knowledge of the properties of those motions. Bayesian models have been developed based on this combination of inference and prior knowledge, which can account for many aspects of motion perception [95, 99, 100, 104]. These models typically make two important assumptions: that local image measurements are noisy and that velocities tend to be slow. These assumptions have been formulated, in the context of the models, in terms of probability distributions.

The assumption that measurements of image motions are noisy, has been instantiated in terms of a specific noise model for how the measurements are made. A standard choice is to assume that the measurements are contaminated by addi-

tive, independent, Gaussian noise with a known standard deviation [95, 99, 100]. For a single moving 1D pattern, there is a family of possible velocities that lie along a constraint line, and the measurements fall along a constraint line, which is 'fuzzy' as a result of the added noise. The 'fuzziness' of the constraint line is related to the uncertainty in the measurements. This relationship is represented as a conditional probability of the measurements given the actual velocities. The noise model constrains a likelihood function, which is a conditional density over the space of 2D velocities, for a particular measurement. The assumption that velocities are typically slow has been supported by psychophysical measurements, and is instantiated by a prior distribution that favors slower speeds [104]. This prior distribution specifies the probability of encountering stimuli moving at any specific speed.

My model is related to these, in that the initial stages involve estimating the 2D velocities of the component gratings, and combining those estimates with a prior for velocity. Like them, my model includes measurements of velocities in the world, which in my case are estimates of the normal velocities of the component gratings of a plaid. I have included not only noise in the measurement of speed, but also in direction, which corresponds to orientation in 2D velocity space. My model also includes a likelihood function which relates the noisy measurements to possible pattern velocities that are consistent with them, which I will describe below. This function is a conditional probability function that is considered as a function of the velocities with the measurements fixed. Having said this, my model is not about estimating motion so much as selecting between two different interpretations of the stimulus. The interpretations are that the plaid is moving coherently or that the plaid is moving transparently.

2.2 Methods

2.2.1 Psychophysical experiments

Four male subjects with normal or corrected-to-normal vision participated in my psychophysical experiments. Experimental procedures were approved by the human subjects committee of New York University and all subjects signed an approved consent form. Three of the subjects (1-3) were not aware of the purpose of the study.

Subjects were presented a centered circular patch containing an upward drifting square-wave plaid for 1.5 s. Randomizing the direction of the plaid, instead of presenting it only in the upward direction, would likely have minimized the possibility of any influence of directional adaptation on the precept of the plaid. This would have also likely minimized any possible influence of preferences to see cardinal directions of motion [139]. The circular stimulus patch was 5 deg in diameter and included a circular fixation mark composed of a black circle within a larger white circle. The spatial frequencies of the component gratings were 1.5 c/deg. Their contrast was 0.4. The stimulus was spatially windowed with a cosine function and temporally windowed with a raised cosine. The temporal windowing increased the contrast to its maximum value by 1% of stimulus duration, 15 ms after onset (in approximately two frames). After the presentation of a plaid, only the central fixation mark was presented. While continuing to fixate during this period, subjects indicated whether the plaid appeared to move coherently or transparently by pressing a key. The response period lasted 1 s and was followed by a 1 s blank period. Brief sounds were presented at the onsets of the test period with the drifting plaid and the response period to help subjects time their responses.

My instructions to the subjects at the beginning of a block trials were brief and straightforward. I reviewed the sequence of a single trial, by showing them a slide of the task. I explained to them that the task consists of three periods: for the presentation of the plaid, for a response to it, and for a brief blank period before the start of another trial. The three naive subjects were generally aware that plaids can appear to move in two different ways, but were unfamiliar with how any particular parametric manipulation would influence the relative dominance of either percept. If needed, I would allow them to practice the task for a minute or two, during which they would experience the two percepts. The task seemed to be easy for them to learn, and they picked it up quickly. I would restart the task immediately thereafter to collect the data I have described here.

I presented plaids with component speeds from 0.5 to 5 deg/s in 0.5 deg/s increments. Their pattern speeds were from 0.5 to 5 deg/s faster than a given component speed in 0.5 deg/s increments. For example, plaids with the slowest component speed of 0.5 deg/s were presented at pattern speeds from 1 to 5.5 deg/s. The angle between the directions of the component motions was set according to the relationship between the component and pattern speeds. For plaids with a component speed of 0.5 deg/s, the half-angles between the directions of the component motions ranged from 60 to 84.8 deg. For plaids with a component speed of 5 deg/s, the half-angles ranged from 24.6 to 60 deg. The sequence of trials was randomized within the ten by ten matrix of speeds. At least six trials per condition were collected. Some subjects completed a small number of practice trials prior to the start of a block of trials to become familiar with the task.

2.2.2 Simulated percept for a single trial

I created an ideal observer model to simulate a subject's percept as formed by an optimal statistical inference process. I defined $H \in [coh,\ trans]$ to be a discrete variable that represents either the *coherent* or *transparent* percept and $\vec{m}$ to be the noisy measurements of the visual information that an observer has access to. For a single trial, an ideal observer makes a decision $\hat{H}$ that is the most probable interpretation (MAP) as a function of the measurement $\vec{m}$, thus

$$\hat{H}(\vec{m}) = \text{argmax}_{H \in \{coh,\ trans\}} p(H|\vec{m}) \tag{2.1}$$

where the conditional probabilities under each condition are computed as

$$p(H_{\text{coh}}|\vec{m}_1, \vec{m}_2) \quad \propto \quad p(H_{\text{coh}}) \int p(\vec{m}_1, \vec{m}_2|\vec{v}) p(\vec{v}) d\vec{v} \tag{2.2}$$

and

$$p(H_{\text{tran}}|\vec{m}_1, \vec{m}_2) \quad \propto \quad p(H_{\text{tran}}) \int p(\vec{m}_1|\vec{v}_1) p(\vec{v}_1) d\vec{v}_1 \int p(\vec{m}_2|\vec{v}_2) p(\vec{v}_2) d\vec{v}_2 \tag{2.3}$$

using Bayes' rule. The first (coherent) term contains an integral over all possible 2D velocities $\vec{v}$ that are consistent with normal velocity measurements arising from both gratings. The second (transparent) term contains integrals over all possible 2D velocities consistent with the normal velocity measurements of each grating individually, and these are then multiplied (since they represent independent probabilities). In order to simulate the model, one determines the form of the likelihood (e.g. $p(\vec{m}_1|\vec{v}_1)$), as well as the form of the prior over velocity $p(\vec{v})$ (which is assumed to be the same for $\vec{v}_1$ and $\vec{v}_2$), and the (scalar-valued) prior on

coherence, $p(H)$.

Note that 2.2 will depend on two pieces of 'prior' information: the probability of coherence, and the prior over velocity. Also, note that the m's are measurements of the normal velocity of each grating. The v's are always 2D translational motions in the image plane. I did not include a normalization constant because the decision is computed by maximizing over H, and the normalization is only a function of $\vec{m}$. I calculated the perceptual decision from the ratio of the two posteriors.

2.2.3 Likelihood function for 2D velocity

To formulate the likelihood, $p(\vec{m}|H)$, I defined what visual information $\vec{m}$ the observer uses to make the decision and how this information is corrupted by noise along the visual processing cascade. I assumed that $\vec{m}$ depended on the experimental setup and the stimulus parameters and wrote the pair $\vec{m} = [\vec{m}_1; \vec{m}_2]$ to represent the measurements of the two component normal motions. The components had fixed spatial frequencies, and conjointly varied only in velocity, which I wrote as $\vec{v}_{c1}$ and $\vec{v}_{c2}$, respectively. I assumed that on each trial, the observer made a measurement of these two component vectors. This assumption allowed me to formulate the model in a compact way and to constrain the likelihood, $p(\vec{m}_i|\vec{v}_i)$, by using the characterizations of previous studies [104]. $\vec{v}$ represents the 2D velocities of a given component grating, which can include components parallel to the grating's orientation.

I defined the characteristics of the measurement for a single set of component velocities. This measurement is an estimate of the normal velocity, and is related to the orientation and speed of the drifting gratings. I assumed that the uncertainty in the measurement of speed was proportional to speed [104]. I also assumed that

uncertainty in the measurement of direction is constant with speed and roughly one third of the magnitude of the uncertainty in speed [43]. Although speed and direction have different units, this ratio can be implemented as the uncertainty in direction being in the orientation perpendicular to the orientation of the vector for the measurement's speed. From these assumptions, I approximated the conditional probability $p(\vec{m}_1|\vec{v}_c)$ as a two-dimensional Gaussian distribution with standard deviations that scale proportionally with component speed, but remain in a ratio of one third. I formulated this elliptical Gaussian in Cartesian coordinates as

$$\begin{aligned} p(\vec{m}_1|\vec{v}_{c1}) &= \frac{1}{2\pi\sigma_v\sigma_\theta} \times \\ &\exp \Big(-[r(v_{c1;x} - m_{1;x})^2 \\ &+ 2s(v_{c1;x} - m_{1;x})(v_{c1;y} - m_{1;y}) \\ &+ t(v_{c1;y} - m_{1;y})^2]\Big) \end{aligned} \tag{2.4}$$

where σ_v and σ_θ are the standard deviation in component and angular (perpendicular) direction, respectively. The parameters r, s, t are necessary for the trigonometric transformation. They are

$$r = \frac{\cos^2\alpha}{2\sigma_\theta^2} + \frac{\sin^2\alpha}{2\sigma_v^2} \;;\; s = -\frac{\sin 2\alpha}{4\sigma_\theta^2} + \frac{\sin 2\alpha}{4\sigma_v^2} \;;\; t = \frac{\sin^2\alpha}{2\sigma_\theta^2} + \frac{\cos^2\alpha}{2\sigma_v^2} \tag{2.5}$$

with $\alpha = \arctan\frac{v_{c;x}}{v_{c;y}}$ being the angle of the grating to the horizontal. 2.4 provides a single measurement of normal velocity of one of the two component gratings. It is drawn from the 2D Gaussian that I have described. Each measurement has a corresponding constraint line, the next step is to work backwards from this, to say something about the 2D velocities that could have given rise to such a

measurement.

While 2.4 describes the forward conditional probability of the measurements given the stimuli (i.e. the component velocities of the gratings), the formulation of the likelihood (the backward or inference direction) needs some additional definition due to the aperture problem. Although the measurements of the component velocities are uniquely defined, their inverses are not. I, therefore, integrated the likelihood as defined in 2.4 over all possible component velocities to compute it for any possible stimulus velocity. To be clear, I integrated over component normal velocities that are consistent with a proposed 2D velocity. The locus of such component normal velocities is a circle with a diameter from the origin to the 2D velocity [7]. This circle comes from the geometry of the intersecting constraint lines. The speeds of the normal velocities are a function of the cosine of the difference between the direction of 2D velocity and the normal velocities. This is done for each 2D velocity. For the numerical implementation of this integral, it was important to integrate over identical segments along the circle, equivalent to equal increments of the angle ϕ. ϕ is a variable of integration, for computing the integral along the circle. I parameterized the possible component velocities and integrated over ϕ,

$$p(\vec{m}|\vec{v}) = \oint p(\vec{m}|\vec{v}_c(\phi))\ d\phi\ . \tag{2.6}$$

For a given 2D velocity, there is a family of possible normal velocities that are consistent with that 2D velocity [7, 95]. These normal velocities lie along a circle that passes through the origin and through the point of the 2D velocity. The speed of the normal motions is a function of the cosine of the difference between

the direction of the normal velocity and the 2D velocity. In addition, there is uncertainty about the exact normal velocities, and therefore, the circle has some width. In particular, this uncertainty has been shown to increase with speed [104]. As a result, the probability mass at points on the circle that are further away from the origin is lower than at points near the origin.

Given a particular measurement there are velocities along a constraint line that could have produced that measurement. For each point along this constraint line, which is to say for each velocity, there is a circle that passes through the origin, through the point of the measurement, and through the point for the pattern velocity. For pattern velocities just flanking the measurement, the circle is comparatively small. Similarly, for pattern velocities further away from the measurement, and extending out along the constraint line, the circle gets larger. This means the probability mass is distributed out over a larger area, and is comparatively flatter and lower than the probability mass distribution for the smaller circles. This would mean that the probability of the pattern velocities decreases extending out along the constraint line away from the measurement. But there is a competing influence. At faster pattern velocities, the circle is not only larger, but it is rotated in such a way that there is a higher peaked mass of probability at the measurement. A consequence of these competing influences is that the velocities flanking the measurement tend to be more likely to have given rise to the measurement.

Having defined the likelihood, $p(\vec{m}_{1,2}|\vec{v})$, of an arbitrary stimulus velocity causing the measurement of velocity for each of the components, I formulated the likelihood for a coherent or transparent percept for a given plaid. The coherent interpretation is consistent with both components having the same rigid velocity

$\vec{v}$, thus

$$p(\vec{m}|H = coh) = \int p(\vec{m}_1|\vec{v})p(\vec{m}_2|\vec{v})p(\vec{v})\ d\vec{v}\ . \tag{2.7}$$

Conversely, the transparent interpretation is consistent with both components having independent velocities, thus

$$p(\vec{m}|H = tran) = \int p(\vec{m}_1|\vec{v}_1)p(\vec{v}_1)\ d\vec{v}_1 \int p(\vec{m}_2|\vec{v}_2)p(\vec{v}_2)\ d\vec{v}_2\ . \tag{2.8}$$

2.2.4 Prior distributions

I assumed that the prior densities, $p(\vec{v})$, that I included in 2.8 to formulate the likelihoods for the two interpretations were identical. I used a heavy-tail power-law density that falls off with speed [104]. I assumed that the shape of this prior did not depend on direction, that it was circularly-symmetric in 2D velocity space. To this point, I had defined the main components of my ideal observer model. The only undefined quantity was the prior on the interpretation, $p(H)$. This prior probability on interpretation was a free parameter. It was not known what the value of this prior should be, although it should be possible to derive it from an analysis of natural images. This prior on the interpretation may affect the measurements, although there is no known results to suggest it does. Without any evidence constraining how they relate to each other I assumed separability.

2.2.5 Simulated percept for many trials

The model did not provide a closed-form solution to the percentage of coherent versus transparent percepts for a given stimulus configuration. It generated an

interpretation based on a single set of noisy measurements $\vec{m}$. To simulate the psychophysical experiments that my subjects completed over several trials of each stimulus configuration I did the following: generated the measurements $\vec{m}$ from the component velocities and the noise model; computed the posterior from the likelihoods and priors; generated $\hat{H}$ based on the posterior-ratio; repeated this many times in order to have sufficient statistical power.

2.2.6 Comparison to other models

I compared the model to two previously described explanations for the effects of speed and angle on plaid perception. I used maximum likelihood estimation to determine parameters for a set of descriptive models that would best explain the psychophysical data I collected.

Following the simplified model of plaid perception described by Farid and Simoncelli, I considered a model in which the pattern of percepts consists of three regions in component and pattern speed space [20, 21]. In this 'speed' model, transparency dominates for pattern speeds faster than, and component speeds slower than, a critical speed. I computed three probabilities for each of the three different regions in the space of speeds that I explored. I computed the likelihood of the psychophysical data for all possible arrangements of the three (or in some cases two) regions as determined by varying the critical speed from slightly slower than 0.5 deg/s to slightly faster than 10 deg/s. I chose the arrangement from a single critical speed for each subject that best explained the data.

I also simulated a model in which the percept is influenced only by the plaid angle and is more transparent at broader angles ('angle' model). I fit a 4-parameter psychometric function to the perceptual data for each subject after collapsing it

onto plaid angle. I used the psignif toolbox (version 2.5.6) for Matlab, which implements the maximum-likelihood method described by Wichman and Hill [140]. I used the probabilities from each fitted psychometric curve to predict the probability of the percept at each combination of component and pattern speed.

I compared the likelihood of the data under three different models, along with two 'bounding' models. I used maximum likelihood to estimate a single probability that best explained the pattern of percepts for all combinations of component and pattern speeds ('coin-flipping' model) as well as a set of such probabilities, one for each combination of speeds ('omniscient' model). I used the values of the likelihood of the data from these coin-flipping and omniscient models as estimates of the upper and lower bounds the likelihood of the data given the output of a model. I normalized the values of this metric from the other models to these bounds.

I used the mean log-likelihood for the data as a measure of the goodness of fit. For this, I computed the probability of observing a particular set of percepts for repeated trials for a single combination of component and pattern speed given a probability from one of the models. I computed the binomial probability density function (pdf) for observing n transparent responses in N trials for a given p. In cases where the numerical value of this pdf was negative infinity, I replaced the value of p with the expected value of $p|n, N$ assuming a flat prior. I recomputed the pdf using that alternative estimate of p. I computed the mean of the log of a set of 100 probabilities from all speed combinations that I measured in my psychophysical experiments.

2.2.7 Determining parameter values

I optimized the values of a few parameters of the model such that its output would closely match the psychophysical data. I could not use iterative optimization methods, such as gradient descent, to determine the optimal values because of the computational cost involved in doing so. I instead simulated the model at discrete sample grids in which all possible combinations of a few values for a small number of parameters were tried. In general, I confined this exploration to be within the extreme values of each parameter as estimated from previous studies [104].

I focused on three parameters: the slope of the speed prior, the limit of the speed prior at which the slope decreases for faster speeds, and the prior on interpretation. Two parameters for the speed prior had been estimated with a different set of subjects [104]. I chose values for the prior on interpretation so that the model appeared to produce reasonable results. The model included five other relevant parameters related to the 2D likelihood. I initially fixed four of these to be the mean values from the previous estimates. The other was the ratio of uncertainty in speed and direction, which I assumed to be 3:1.

The first grid that I ran was for a combination of $5 \times 5 \times 5$ values for the three parameters of the two prior distributions. In general, the maxima of the likelihood surface seemed well centered and smoothly varying within this three-dimensional space. Nonetheless, I wanted to check for other maxima at locations within the $5 \times 5 \times 5$ grid. I computed interpolated likelihoods for a $105 \times 105 \times 105$ grid. For two of the four subjects, this produced slightly higher likelihoods after simulating the model with parameter values near maximal in the interpolated set. At that point, I chose whatever parameter values produced the highest likelihood, either from the interpolated estimates or from the original $5 \times 5 \times 5$ grid.

I ran one additional simulation to determine whether different combinations of the values of the likelihood parameters would allow for a better fit. I had not explored these parameters up to that point. I varied four of the five parameters with values in a $3 \times 3 \times 3 \times 3$ grid. I used values for the parameters of the two prior distributions that gave maximal likelihoods for each subject as I had estimated them through the previous explorations. I used the best pattern of simulated responses from this grid for comparing to the other models.

2.3 Results

2.3.1 The effects of component and pattern speed

I have plotted my psychophysical and model results within a common space of component versus pattern speed. Figure 2.2b shows examples of the plaid stimuli represented in different regions of this space, as well as vectors for each whose lengths correspond to the example plaid's relevant velocities. Slow plaids are represented in the lower left, whereas faster ones are represented in the upper right. Moving upward in this space corresponds to increasing the pattern speed, increasing the plaid angle while holding the component speeds fixed. Extending radially away from the origin along a line corresponds to maintaining the ratio of component to pattern speed such that the plaid angle remains constant. I did not sample in terms of fixed plaid angles, but in this space they increase radially in a counterclockwise direction away from the abscissa. The percepts for each subject, for each combination of the two speeds, corresponds to the brightness within each of the the box-shaped points shown in Figure 2.2b. The brightness relates to the ratio of the two percepts across trials (black = 100% coherent, white = 100% transparent).

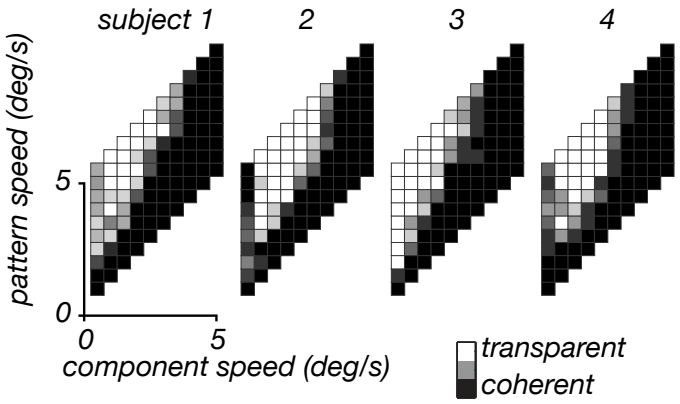

Figure 2.3: Psychophysical results from four subjects with drifting square-wave plaids at 100 different combinations of component and pattern speed. The whiteness within each box represents the ratio of the two percepts across trials. For all subjects, plaids appeared coherent for an intermediate range of component speeds when the plaid's pattern speed was slightly faster than its component speed.

The percept of a plaid was strongly influenced by its component and pattern speeds (Figure 2.3). In that it covaries with the relationship between these speeds, the angle between the directions of the component motions affects the percept. At slow component speeds, transparency increased as pattern speed increased. Plaids with pattern speeds faster than the pattern speed at which the percept transitions from coherent to transparent appeared transparent through the fastest pattern speeds. At the slowest component and pattern speeds, plaids appeared coherent. At the fastest component speeds, faster than about 3 deg/s, plaids appeared coherent through the highest pattern speeds. This corresponds to the upper rightmost set of points, where there is a transition in the plots from left to right from transparent to coherent. At the slowest component speed, 0.5 deg/s, some subjects perceived plaids to move coherently for all pattern speeds, whereas one saw plaids at these speeds to always be drifting transparently. This coherence at slow component speeds corresponds to a ridge of darkened points at the leftmost

edge of the plots within Figure 2.3. The shape of the transparent region varies across subjects. For subject 3, for example, it is more angular through the range of intermediate component speeds at which plaids were transparent for some pattern speeds. In general, there was not a specific angle, or pattern speed, beyond which transparency occurred for all subjects.

2.3.2 Bayesian observer model

The visual system of human observers has a preference to 'see' slower speeds. They are assumed to be more probable [95, 104, 100]. Plaids that are drifting at fast speeds might, therefore, appear transparent because the speeds that are consistent with the coherent percept are assumed to be less likely [20]. This might account for the increased transparency for plaids at many of the combinations of speeds for which I measured the percepts. But my results are inconsistent with this explanation for some sample points. Plaids with fast component speeds, for example, appeared coherent through the fastest pattern speeds. Plaids with these speeds appeared coherent at pattern speeds at which the percepts for plaids with slower component speeds appeared transparent. This could not be accounted for if the coherent percept was simply dependent on level of improbability that results from a faster pattern speed. The increased coherence that I observed may result from a second preference to form singular perceptual interpretations, thus to see one pattern of motion.

I constructed a Bayesian observer model to test this explanation for the effects of speed on plaid perception. Figure 2.4 shows a graphical representation of my model for a single presentation of a drifting plaid. The images within this figure are from a plaid whose speeds are near the center of the parallelogram pattern

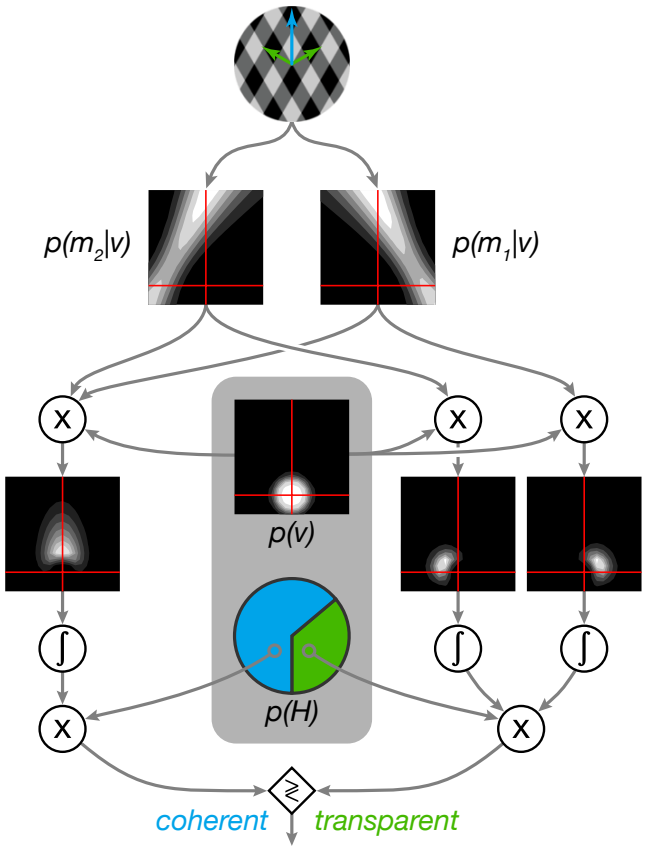

Figure 2.4: Bayesian observer model for plaid perception. All plots with red overlaid axes are for the same region of 2D velocity space. I illustrate the constituents of my observer model for the perception of a single presentation of a drifting plaid with a single component and pattern speed. First an observer forms independent representations of the leftward and rightward component motions. The likelihood functions formed by the components are the probabilities of the internal measurements given any velocity. Two preferences of the visual system are represented within the gray rounded rectangle in the center. The prior for velocity confers a preference to see slow speeds and the lower prior for the hypotheses favors coherent percepts, which are represented by the blue sector. The likelihoods are combined with the prior for speed either together or separately depending on the hypothesis of the percept. The product of the likelihoods and priors for speed are integrated and multiplied by with their respective priors for the hypotheses. The hypothesis whose posterior is greater is the predicted percept.

of samples within the parameter space of the plaid's speeds. The plaid in this example has a component speed of 2.5 deg/s and pattern speed of 5 deg/s.

In the first stage of the model, an observer forms independent internal representations of the velocities of the components. The internal measurements for these representations are imprecise. Stocker and Simoncelli showed that for 1D motion of broadband patterns, the variability in measurements is proportional to a logarithmic function of speed [104]. I used estimates of the widths of distributions of such measurements from that study to define the widths of the distributions of measurements for the 2D motions in the model. In addition, I assumed that the ratio of the major axes of the 2D Gaussian of measurements in velocity space is 3:1 for speed versus direction [43]. In other words, the measurements have greater variance in speed than in direction.

The actual physical motions that generate a particular set of velocity measurements is unconstrained. To determine the distribution of measurements given all possible velocities in a probabilistic sense, I integrated across all possible pattern motions at fixed directional separations. I also assumed that observers have a circularly-symmetric prior for velocity that falls with speed as a power law [95, 104, 100]. As with the likelihood, the shape that I initially assumed for the prior for velocity was based on measurements from the set of psychophysical subjects from that earlier study.

The prior for velocity was combined with the likelihoods jointly and separately in separate pathways for determining the probability of the coherent and transparent percepts. I computed the product of the prior for velocity with both likelihoods for the coherent hypothesis and with each likelihood for the transparent hypothesis. Figure 2.4 shows this product. For the coherent hypothesis for this example,

it is slightly spread out and curved along directions centered on the pattern direction. It falls off towards faster speeds. Also, it is significantly shifted towards slower speeds than those at which the likelihoods for the component motions intersect. The products for each of the components in the pathway for the transparent hypothesis are smaller and are centered near the directions of the component velocities.

These three products are integrated and multiplied by a prior, of two constants, which favors the coherent hypothesis. This is represented in Figure 2.4 by the green and blue pie plot. The sector for the coherent hypothesis (in blue) is larger. At this point in the flow of information, there are two values, a posterior for each hypothesis. These do not sum to 1 since the probability of the measurements was left out since the decision does not depend on it. The percept is set to be the hypothesis for whichever is greater. The percept would be coherent, as it was for the example trial in Figure 2.4, whenever the posterior for that hypothesis is greater, regardless of how much greater it is than the posterior for the other hypothesis.

2.3.3 The prior for velocity and the prior for perceptual interpretation

Having formulated a model for the perception of drifting plaids, I explored how varying a subset of its parameters affects the predicted percepts for plaids drifting at different speeds. I ran a series of simulations to establish how the shape and values of the priors affect the pattern of plaid percepts. Of course any changes that favor faster speeds or the coherent percept directly should increase the amount of coherence (shrink the transparent region). I initially ran all simulations at each of

the sets of likelihood parameters from the other group of three subjects. I found that the model predictions were about the same in terms of their similarity to the psychophysical data regardless of the set of likelihood parameters I tried. Therefore, I chose one set of parameters and used those for all subsequent simulations.

I simulated plaid perception for seven combinations of the parameters for: the falloff speed and slope of the prior for velocity; and value of the prior for coherence (Figure 2.5). The slope and falloff speed were set to the extremes of previous estimates for 1D motion. I chose high and low values of the other prior that gave results that looked reasonable by eye.

Figure 2.5a shows slices of the prior for velocity as well as the values for the prior for coherence for each combination arranged in a hexagon. The simulated percepts are arranged for illustration so that they are more coherent in the rightmost set of plots. The value of each parameter was varied from its extremes along a line in one of the three orientations within the group of plots. The red highlights are for the parameter that was varied in each case. Figure 2.5b shows the corresponding percepts.

Increasing the prior for coherence (from left to right in Figure 2.5b) shrinks the transparent region. Coherence increases from what appears to be all sides of the transparent region. Plaids are more coherent at the fastest pattern speeds, at slower component speeds and at the slowest component speed with high pattern speeds. The simulated percepts for the lowest and highest values of this prior have more and less transparency than the percepts from my psychophysical subjects.

Decreasing the falloff speed of the prior for velocity (from lower-left to upper-right in Figure 2.5b) increases coherence. The changes in the amount of transparency and coherence that results from changing the falloff speed are different

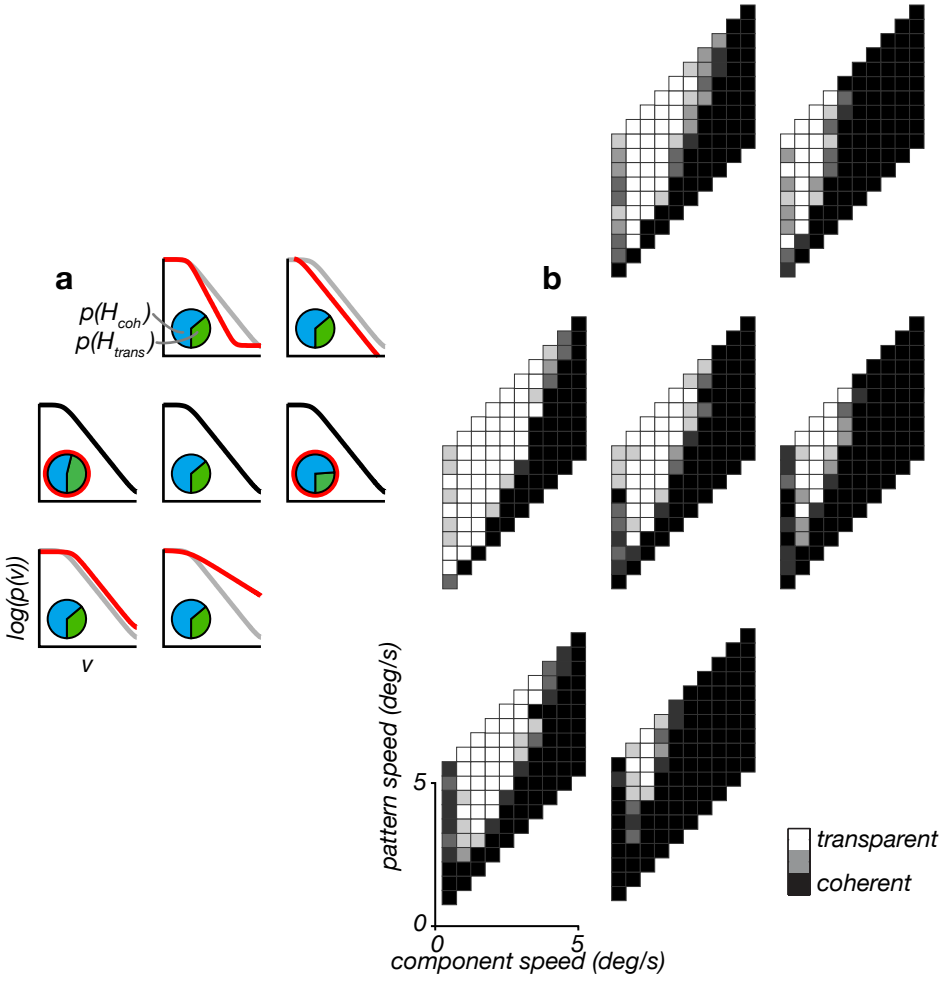

Figure 2.5: The effects of manipulations of the two prior distributions. (a) Each plot shows a cross-section of the prior for velocity. Inset is the pie-shaped plot whose sectors correspond to the values of the prior for the two percepts (blue for coherent, green for transparent). Variations from low to high values for three parameters: the falloff speed and slope of the velocity prior and the value of the prior for coherence are shown. Combinations of these parameters are arranged in a hexagon. The middle set is the default values. The prior for coherence increases from left to right. The falloff speed decreases from lower left to upper right. The slope increases from top to bottom. The red highlights whatever is changed in the priors for a given position. (b) Simulated percepts for the combinations of parameter values in the corresponding positions in (a). The size and shape of the transparent region is affected.

than for changes in the prior for coherence. When the falloff speed extends to faster speeds there is more transparency, but the ridge of coherence at the slowest component speed is increased. Plaids with those speeds are more likely to appear coherent. When the falloff speed is low, there is more coherence, but also less of a ridge. These effects are like a leftward shift in the transparent region for the plots of the plaid percepts.

Increasing the slope of the prior for velocity (from upper-left to lower-right in Figure 2.5b) increases coherence. The magnitude of the ridge of coherence increases, but the main effect is that the transparent region shrinks significantly. Plaids at many combinations of component and pattern speed that are transparent in the default case are coherent when the slope is more positive, and vice versa when it is more negative.

Model predictions with reasonable parameter values for these priors generate percepts that are similar to my psychophysical data. The variability that results through extreme values of these parameters parallels the variability in my psychophysical subjects. Changing the values of these parameters affects the overall pattern of percepts in predictable ways. For example, increasing the slope of the prior for speed such that it favors faster speeds, increases coherence.

2.3.4 Fits of Bayesian and alternative models to data

I optimized for three model parameters of the two priors shown in Figure 2.5a by hand-fitting them with discrete parameter grids and using spline interpolation to estimate optimal parameter values within those grids. By comparing the results of this optimization to other models it was possible to assess how well the model could match the percepts. I included two extremal models for the upper and lower

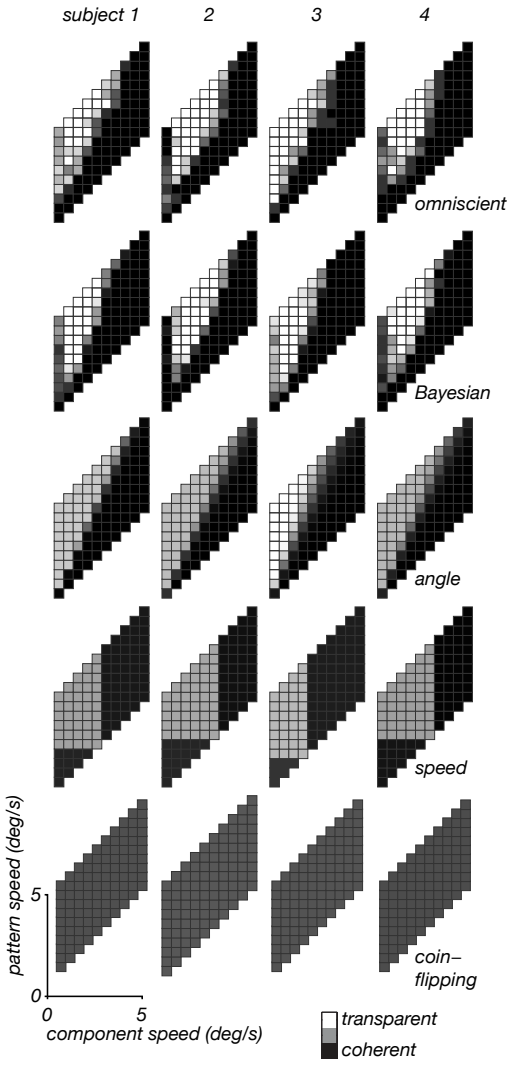

Figure 2.6: All plots are in the standard space of component and pattern speed with variation in the percepts shown in white to black within each point. The predicted percepts from five models of plaid perception were fit for each of the four psychophysical subjects. The omniscient and coin-flipping models in the first and fifth rows were the upper and lower bounds, for which 1 and 100 different p-values were computed by maximum likelihood estimation for the 100 combinations of plaid speeds. The angle model was determined for each subject by fitting a psychometric function to their data when collapsed onto the plaid angle. The speed model was determined by fitting 3 p-values in regions formed by all possible cutoff speeds. The results from the Bayesian model that are shown are from only a few trials. The model was run with parameters of the prior distributions that were optimized by hand.

bounds. For the 'omniscient' model (the upper bound), I calculated a p-value from the psychophysical data for each combination of speeds using maximum likelihood estimation. These 100 p-values essentially matched the probabilities of the two percepts from the trials for each condition for each subject (top row of plots in Figure 2.6). The 'coin-flipping' model (the lower bound) was fit by calculating a single p-value that best explains all of the perceptual data for each subject, which was also estimated by MLE (bottom row of plots in Figure 2.6). The coin-flipping model captures almost none of the features of the real data. It only predicts that there is a biased mixture of the two percepts so that all of the points from this model are neither black nor white, but in between.

The second row of plots from the top in Figure 2.6 shows the results of my optimization for the three parameters. This best fit of the model closely matches the actual pattern of percepts for all subjects. As in the physiology data, plaids appear transparent from 0.5 deg/s to about 3 deg/s for fast pattern speeds and are coherent for other combinations of the two speeds. The model also captures the overall shape of the pattern of percepts for each subject. This includes a ridge of coherence at the slowest component speed whose intensity varies across subjects. The data for subject 3 does not include this ridge, although the model that is best matched for this subject includes it to some extent. The prediction of my model also differs from the percepts for this subject in that it is less dependent on plaid angle. The model predicts a transparent region that is more rounded.

The 'angle' model was based on the hypothesis that plaid perception can be explained in terms of angle only, larger angles being more transparent. The results for the angle model are shown in the third row from the top in Figure 2.6. Not surprisingly, they have angular-shaped transparent regions. Plaids with large an-

gles between their directions of motion, on the left side of the plot, are predicted to be more transparent. For most subjects, this is less transparent than for the psychophysical data. The actual data is fully transparent for many of the plaids with those combinations of speeds that are mixed in the simulated data, and fully coherent for others. The angle model also fails to predict the ridge at the slowest component speeds and the rounding off of the transparent region at the faster component speeds.

The 'speed' model was based on the hypothesis that there is a critical speed at which plaids with faster pattern speeds appear more transparent. The speed model is similar to the angle model in that it does not predict fully transparent or coherent perception for plaids that are one or the other in the real data. The results from the model are gray at several points that are fully white in the real data and for my Bayesian model. The speed model is correct, however, in predicting that plaids are more transparent at faster pattern speeds. It also does not capture the ridge or the rounding at faster component speeds, which my Bayesian model does. Interestingly, Subject 3 has a fit for the speed model that becomes transparent at pattern speeds that are slower than for the other subjects. The real data also do not have abrupt transitions like this model, where the percepts shift from all coherent, for example, to all transparent without samples which are a mixture of these extremes. Smoothing these transitions by, for example, adding parameters to fit that transition would, of course, allow for a better match to the real data. It seems clear, nonetheless, that such a model would not fit the data as closely as my model, which has fewer degrees of freedom.

To summarize, my Bayesian observer model captures many features of the pattern of plaid percepts that neither of the two compared models did. I computed

a metric to capture the goodness of fit for these models, which is explained in the following section.

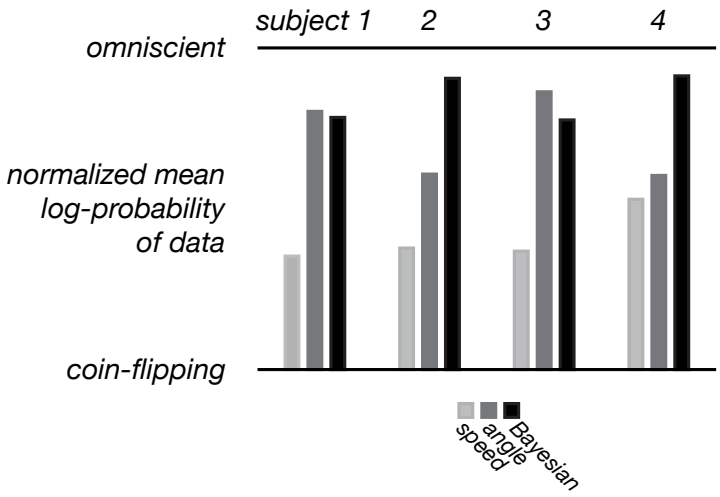

Figure 2.7: Comparison of model performance. Normalized mean log-probability of the psychophysical data for the three comparison models for the four subjects. Probabilities for each subject were normalized to the probabilities from their coin-flipping and omniscient models.

2.3.5 Comparison to alternative models

I quantified the quality of the fits for the models shown in Figure 2.6 by computing the mean log-probability of data for all trials and for all speeds for each subject (Figure 2.7). For subjects 2 and 4, the Bayesian model fits the psychophysical data nearly as well as the omniscient model and better than the other two models. For subjects 1 and 3, the Bayesian model is slightly worse than for the other two subjects, but it still performs well. For these subjects, the angle model performs better than for the other subjects. The speed and angle models have 4 free parameters, which are fit for each subject. The Bayesian model includes several parameters for the likelihood that I did not fit, and my optimization was for only three parameters

of the two priors. In addition, the goodness of fit for the Bayesian model is limited in that the process for estimating optimal parameters was limited.

2.4 Discussion

My psychophysical results suggest that: (1) plaids with slow component speeds and faster pattern speeds appear transparent; (2) plaids with fast component speeds appear coherent, which is also the case for some subjects at very slow component speeds; (3) covarying with these, plaids with greater angles between the directions of the component motions appear transparent, with angle having a greater influence for some subjects.

I have described a probabilistic model for forming a percept of a drifting plaid, linking steps in this process to realistic preferences of human observers, and their mechanisms for inferring motion, when presented with a sequence of moving visual patterns. My modeling results suggest that a Bayesian estimator that includes two priors, which favor slow speeds and singular percepts, can account for the psychophysical effects that I measured. The model consists of the following processing stages: (1) a prior for slow speeds is combined with noisy internal measurements of the motions of the components; (2) the products from step 1, are integrated and combined with a prior that favors the coherent percept; (3) the single values from step 2 are compared and the perceptual hypothesis for which there is greater evidence is chosen.

By simulating the percepts for several trials for the same plaid speeds as in my psychophysical task, I found that the shape of the prior for velocity or the value of the prior for coherence, impacts the pattern of percepts in consistent and

predictable ways. Coherence increases, for example, when the prior for velocity is of higher probability through faster speeds. Coherence likewise increases with greater values for the prior for coherence. I optimized the shape and value of these priors to match my psychophysical data as closely as possible. Inter-subject differences in how plaids appear to move can be accounted for in terms of the way that a subject's perceptual systems favors slower speeds and singular percepts.

2.4.1 Relation to previous results

Most of the earlier results on plaid perception differ from mine in terms of the stimulus properties of the plaid that were explored, as well the task and measurements that were used. This limits the extent to which these results can be directly compared. That said, I would like to make a few observations about the earlier findings and will compare my results to the extent possible.

Farid and Simoncelli's results are similar to mine in terms of the effects of plaid component and pattern speed on plaid perception [20, 21]. They measured percepts for different combinations of component and pattern speed with square-wave plaids. They found that plaids are coherent at slow component and pattern speeds and transparent for faster pattern speeds. Importantly, they suggested a link between the effects of speed on plaid perception with the preference of the system to see slow motions. They recognized that results, such as those by Thompson showing that low contrast gratings appear to move slower than they actually are, apply to the perception of drifting plaids [93].

My results differ from their results in a few ways. I did not find that there is a critical pattern speed beyond which plaids at all component speeds were transparent. This difference may primarily be the result of differences in the speeds at

which I sampled the percepts. Although they sampled at many combinations of speeds, my sampling was more uniformly spaced, which might have made it easier to see that there was not a single cutoff speed. Their data are more coherent at slow component speeds. Their subjects transition to transparent percepts at faster pattern speeds than my subjects. This could also be a result of sampling or that there were different subjects in both cases. I also found that a model based on the model that they suggested could not account for my psychophysical results as well as a Bayesian model.

Earlier studies of plaid perception include those by Adelson and Movshon, who measured the effects of changing the relative contrasts or spatial frequencies of the components [7]. For sine-wave plaids with a fixed plaid angle and with one component at a fixed contrast of 0.3, they measured which percept occurred during brief presentations as the contrast of the other component was varied. They did not provide much detail about these results and did not present them in full, but only summarized their findings. From what they did publish, coherence increased as the contrast of the component with varying contrast was increased, when it was more similar in contrast to the other component. In a second set of experiments, they used the relative contrast required for subjects to perceive coherent motion as a measure of the effect of altering the relative spatial frequencies of the components. The amount of contrast (similarity in contrast) required for coherence increased as the differences in the spatial frequencies increased. So plaids were less coherent when the spatial frequencies of their components were more different.

I did not explore plaids with components that differed in any way besides having different directions of motion. I also used square-wave instead of sine-wave components. I believe that sine-wave plaids are more coherent than square-wave

ones. I do not know why the presence of higher frequency components in the spectra of square-wave components makes them more transparent. I used square-wave components in part because I wanted there to be a sufficient proportion of transparent percepts in the region of the parameter space that I explored. My results would likely be similar regardless of this difference.

My model could generate predictions for the ratio of the two percepts when the components have different spatial frequencies, although I have not explored that. Differences in the components should change the likelihoods. As it is, my model might not be influenced by potentially important characteristics of the stimulus in this case. For example, the number of cycles within the aperture might be an important confound. Components with few cycles within the aperture might significantly increase the variability in estimates of the velocities of the components. This could affect the psychophysical results, but not be there in the model. A lower limit on the number of cycles within the stimulus could be set, although it would be impossible to dissociate its influence from the overall spatial frequency.

More recent work by Smith has focused on the perception of plaids with components that differ in spatial frequency [13]. Subjects viewed drifting horizontal plaids with a constant angle between the directions of the components. The components had equal contrast. For five different overall contrasts, he determined the spatial frequency difference (in octaves) at which coherence was reported on 50% of the trials. This was done for plaids with five different component speeds. For the three subjects, the coherence limit decreased in a roughly linear way for increasing speeds. The limits were shifted towards smaller limits for lower contrasts. The largest ranges of spatial frequencies for which coherence occurred on half of the trials were for slower component speeds and higher contrasts.

My results are consistent with this, in that plaids were less coherent at faster component speeds, although in my data this was true within a limited range. I did not explore the effect of overall contrast, although I would expect lower contrasts to favor the transparent percept, in contrast to Smith's results. The likelihoods in my model should broaden from increased internal noise at low contrasts. Therefore, the posteriors should be more influenced by the prior for slow speeds, which favors transparency.

Unlike Smith, I did not observe greater variability in the percepts for plaids at correspondingly narrow plaid angles. Smith used plaids with a half angle of 45 deg, less than most of the plaids that I explored. Plaids with similar angles are represented in the the lower right region in my plots in the space of component versus pattern speed. In my data, subjects consistently saw the coherent percept through a similar range of component speeds. It is unclear what accounts for this disparity. There may be other indicators in my model that would reflect increased variability at narrow angles, such as differences in the posteriors. Variability in my posteriors would not show up in terms of the probability of either percept across trials, which is what I considered.

Studies that have focused on plaids for which the components are different in parameters other than direction, such as the three I discussed above, have consistently confounded the effects of the relative differences and the overall values of a given parameter. For example, coherence may be more likely when the contrast of the components are more similar, which is Adelson and Movshon's interpretation, or more likely when the overall contrast is greater, these not being mutually exclusive [7].

Another suggestions is that plaid angle has a greater influence on plaid per-

ception than other parameters, including differences in the contrasts and spatial frequencies of the components [138, 14]. Kim and Wilson showed that angle has an influence, although their results are difficult to compare to since they measured the probability of coherence for plaids with components that had very different spatial frequencies and for only five angles at a one component speed. They did not include a control in which the components had the same spatial frequencies. They did not consider that their results might have been different at other speeds. As I have pointed out, changing the angle corresponds to changing pattern speed and their data could not differentiate between the effects of these parameters.

A plaid's angle impacts the percept at some component speeds, including at the component speed that Kim and Wilson explored [14]. Plaids are more transparent at broader angles (at higher pattern speeds), consistent with their results. But at the fastest component speeds in my data, angle did not have a significant effect; plaids appeared coherent at those speeds for many different angles. My modeling results suggest that plaid perception cannot be fully accounted for in terms of only the effects of angle.

Differences between the above-mentioned studies further support the view that the effects of a plaid's parameters on how it appears to move depend on the parameter ranges within which any manipulations are explored. Claiming that any parameter influences the percept in a particular way or that one manipulation has a greater influence than any other is of lesser utility in this light. I do not suggest that the relationships between speeds and percepts that I observed would extend to all other speeds or would hold for very different stimulus configurations. I could not conclusively predict, for example, what the percepts would be at different contrasts or for different spatial frequencies. One of my motivations in grounding any

psychophysical data to a model of how the perception is formed is to solidify the link between a broader region of the space and the perception of integrated or segmented motion. Models will help to inform how far within any region of the space any relationship should exit. Moreover, my model also facilitates linking the psychophysics to the mechanisms that underlie it, which allows for extensions of this sort.

2.4.2 How to estimate plaid percepts

I want to point to some of the drawbacks of tasks in which subjects indicate whether a plaid appears coherent or transparent. In doing so, I am not suggesting that better alternatives exist. One issue is that data from them depend on subjects making consistent judgements. Variability across trials, as might result from adaptation, including adaptations at different timescales, would make the data more noisy. In my case, this variability would be lost in my measure of the ratio of percepts across all trials. In performing my task, I did not notice any consistent variability across trials of the likelihood of either percept, nor did any subjects describe any such effects. But it is difficult to rule out, and it would be difficult to detect, possibilities such as an increased probability of seeing the other percept after having seen the alternative percept on many previous trials. I know of no reason why this influence, if it exists, would not impact the pattern of percepts equally across speeds, which are presented in a random sequence. It should be relatively straightforward to check for these sequence effects. A better understanding of them would be useful and they are interesting in their own right.

2AFC tasks additionally depend on there only being two possible percepts. If subjects only saw one of the two drifting gratings, for example, would they have

indicated that their percept was transparent, which in other cases would have consisted of a clear percept of two gratings sliding over each other? I assumed that whenever subjects see the percept consistent with the segmentation of the motions, they also see them to move transparently.

Another source of variability could result from bistability [138]. The percept could switch from one percept to the other, or to some other percept, during the brief presentations that I used. I believe that this possibility is minimized by the briefness of my presentations, which were significantly shorter than the average durations to transition that have been previously reported. Those estimates are roughly an order of magnitude longer than my presentation of a plaid in a trial. I do not know how subjects would deal with transitions were they to occur. Might subjects, for example, consistently pick one of the percepts, such as coherent, if they perceive it at any point during a presentation? Might they instead choose whichever occurred last during the presentation, or up to the point just before they indicated their percept? They might apply some mixture of these sorts of rules or biases. A similar set of uncertainties exist in terms of subjects performing the task. Maybe they were not paying attention, or they simply pressed the wrong key. These effects could be looked for by repeating my task and having subjects continuously report their percepts.

A plaids speed could also be too fast to form a clear percept of whether it was coherent or transparent. Were that to occur would subjects indicate that they saw coherent motion even though they did not? Farid and Simoncelli included unclear percepts in a region of the speed space where the component speed exceeded about 5 deg/s, which was the upper limit for my data [20]. I intentionally confined the speeds of the plaids that I explored to reduce the possibility of other percepts. To

define the region of speeds that I sampled, I ran a series of preliminary experiments with one subject in which a third response in the task was included for indicating that the motion was neither coherent or transparent. I found no such percepts with one subject within or near the region of the parameter space that I considered in this paper.

Subjective categorical judgements of plaids with components that differ in spatial frequency may be especially complicated. When the components are different in that way, there are fewer cues in agreement for the coherent percept [13]. The plaid may appear to be of crossed gratings regardless of whether it appears to move coherently, unlike when the components have matched spatial frequencies. In addition, the plaid direction in the coherent case may perceptually differ from the IOC direction, since the perceived velocities of each component should differ when they have different spatial frequencies.

Alternative methods for measuring plaid perception address some of these limitations, although they may include other drawbacks. Using the bistability measures developed by Hupe and Rubin, for example, would include many of the same problems with choosing a singular percept, while adding significant time to a set of measurements [138]. This would likely require a significant reduction of the number of parameter combinations that could be explored.

2.4.3 Connection to other models of visual motion perception

Several Bayesian models for the perception of visual motion have been described [141, 95, 104, 96, 99, 100]. These models can accurately predict the perception of increasingly complicated and realistic patterns of visual motion. The first part

of my model, which includes a prior for velocity, is an extension of these models. Within the Bayesian framework, priors for velocity allow for the prediction that objects appear to move more slowly at lower contrasts. They suggest further that the likelihood has greater variance at low contrasts from increased internal noise, so that the posterior on which the percept depends is effectively more shifted towards slow speeds.

The shape of the prior for speed was most precisely defined in a recent study by Stocker and Simoncelli [104]. They were able to estimate the shape of the prior from their psychophysical data. My model extends this prior for slow speeds, by assuming that it is circularly symmetric in 2D velocity space and that it is similar for a different set of stimuli and for a different task. I have left out more sophisticated possible features of this prior, such as radial extensions through higher speeds along cardinal and intermediate directions. These would be star-like ridges, which may explain biases towards those directions and orientations, consistent with the 'oblique effect'. I likewise did not include a concentric ring of higher probability at very fast speeds, which might account for recent results that suggest that perceived speed can be biased towards faster speeds for fast motions at low contrasts [142]. The inclusion of these rings may be of minimal importance to my results, where the component and pattern speeds are comparatively slow.

The values for the priors and likelihoods that I used as a starting point for my simulations were measured for a different set of subjects from the earlier Stocker and Simoncelli study [104]. Obviously, my results would be improved by measuring these distributions for the same subjects who complete my plaid task and showing that they produce a close match when used in my model. These measured values could be compared to the ones I found through optimization. Acquiring

these measurements would be an opportunity, though, to extend the method for estimating the priors and likelihoods for motion from 1D to 2D. This could be done in sequential forced choice tasks in which the perceived speed and direction is reported. Including these updated priors and likelihood would allow for further extensions of the model to other aspects of visual motion perception.

My likelihoods were bow tie in shape in velocity space, which is to say that they are wider at directions further away from their normal velocities. The likelihoods in Weiss and Fleet are similar to mine in that respect [99]. The shape of their likelihoods resulted from the inclusion of noise in all derivatives (spatial and temporal) in the representation of a moving edge. Mine have a bow tie shape because I assume uncertainty in the measurements in speed and direction. I assumed that the distribution of measurements had a specific aspect ratio for the elliptical distributions and that that ratio was constant across speeds. This assumes that subjects' measurements are more variable in speed than in direction. In early simulations, I explored likelihoods with very small amounts of uncertainty, so that there was essentially a single narrow line along the constraint line for each component. In that case, my model still produced reasonable patterns of percepts, not entirely dissimilar to those in my final fits to each subject.

I noticed in plots of the final likelihoods that I used that they have higher probabilities in directions flanking the normal velocities for the components. The measurements, in other words, are higher for velocities not orthogonal to the orientations of the component motions. As previously explained, this results from the interplay two influences on how the possible pattern velocities along the constraint line relate to a given measurement. The probability mass along the circle of possible pattern velocities for a given measurement is overall lower, since it is

more spread out. But at the points near the measurement the probability mass has a higher peak, since it is more concentrated at those velocities.

My model condenses all of the processing involved in forming separate likelihoods into a single step. It would be more realistic to include a fuller representation of that process, to expand out the steps involved. A reasonable starting point could be with the formation of pixelated representations of the image sequence, akin to the processing done by the retina. One could then extend from these early representations through cascades of processing capable of representing the pattern motions. This could include a realistic means of parsing the representations of the leftward and rightward components, assuming that they are represented separately.

My model depends on these independent likelihood representations. They are of particular importance to the pathway leading to the transparent percept, where they are separately combined with the prior for velocity. One could imagine a physiological correlate of those representations as component direction-selective cells in striate and extrastriate areas. At the angular separations that I explored, there should be minimal overlap in the responses of these populations to the two components. That said, my preference would be to have a realistic front-end even in cases where the angular separation is much narrower. I believe that those plaids would be represented by the same process, which would need a cleaner way of capturing their interaction.

A related issue for the likelihoods is directional repulsion. The perception of the velocity of one component is affected by the presence of one or more additional components. Their perceived directions may be pushed away from each other, towards greater separations than for the real motions [143, 144, 145, 78, 146, 147]. These repulsive effects are minimal beyond about 40 deg. The inclusion of these

repulsive effects, like an expanded front-end, would improve my model, perhaps allowing for better predictions of plaid percepts over a wider range of stimulus parameters. At the same time, these additions would make my model significantly more complicated. And I believe that additional insight into these and other perceptual effects is needed before building them in is warranted.

The products of the likelihoods with the prior for velocity form intermediate distributions that show significant shifts towards slower speeds for the component and pattern velocities. It would be interesting to know whether the perception of these motions is similar to these distributions. In the subsequent steps of my model, these products are integrated and multiplied by the prior for the percepts, which is greater for the coherent percept. This bias to see singular motions was mentioned by Hildreth, but further work is needed to confirm it [148]. It is possible that experience of motions in the world reinforces that the motion of singular rigid objects is more common than transparently overlaid ones moving in different directions, which seems reasonable. But is the preference only to see one instead of two motions or does it extend to higher numbers of motions, such as four instead of six?

Before optimizing my model to match my psychophysical data, I determined how variations in the two priors impact the simulated percepts. I summarized these findings above, but would like to make a few additional observations. Extending the flat portion of the prior for velocity at the slowest speeds increases transparency, as I have explained. In contrast, increasing the slope of the prior for velocity increases coherence. Although manipulations of either of these characteristics can increase coherence, they show opposite effects at slow component speeds. When the falloff speed is slower, coherence decreases at the slowest component speed

while the overall amount of coherence increases. Conversely, when the slope is increased coherence increases at the slowest component speed while the overall amount of coherence increases. An explanation is that slower falloff speeds favor the transparent percept by shifting the prior to slower speeds when the component speed is slow. At faster component speeds, though, the effect is different because the likelihoods consistent with the coherent percept are greater and there is not a sufficient difference between the component speeds to overcome it. Higher slopes instead simply boost the coherent percept, by increasing the weighting for faster velocities.

2.4.4 Model performance and connection to underlying mechanisms

To determine the goodness of fit of my model, and the comparison models, I computed a metric related to the log-likelihood of the data given the number of trials and the p-values from the model. The first step was to determine the binomial probability density of observing some number of transparent trials given the total number of data trials and the probability of transparency as estimated from the model. In most cases, including for all of the models besides my model, I simply computed the mean of the log of those pdfs to get a single value of the goodness of the fit.

For my model, however, I included an additional step, when necessary, which was to replace the p-values to avoid infinite log probabilities of observing the data. Since I estimated the p-values from the model by generating a sequence of trials and determining the ratio of transparency across those trials it was possible to get a probability of transparency that was 0 or 1. Of course the possibility of

either of these extremes is minimized, but not eliminated, by generating more simulated trials. But there were still cases in which this occurred for the trials that I simulated. If there was a 0 or 1 p-value from the model at a pair of speeds and any trials in the actual data for the other percept, the likelihood would be infinite. An example of this is when the model predicts 0 probability of transparent and there is one of five trials that is transparent. It was important not to simply remove data at these points, which would have given an unfair advantage to model predictions in which this frequently occurred.

So only for the subsets of trials for which the p-value was 0 or 1 and the ratio of observed percepts across trials was not 0 or 1, I replaced the p-value with a different estimate. The replacement p was calculated as the number of transparent trials from the model plus 1 divided by the number of model trials plus 2. For 0 transparent percepts in 10 trials, the replacement p would be 0.083. For 10 of 10, 0.917. Using the replacement p-value I computed the binomial pdf as before and the mean of the log of that. For plots of several models, I normalized and flipped these, which were more negative the worse the fit, by the values from the lower and upper bound models.

One of my subjects did not show any coherence at the slowest component speed. This meant, in comparison to the other subjects, that the angle model did better because it predicts transparency at large angles. I do not know whether this subject would have shown a ridge of coherence at a component speed slower than 0.5 deg/s. It is possible that, in that case, my model would have been able to have a closer fit, since it usually included a ridge of coherence, however slightly.

This raises a general point, which is that the extent to which any of these models can fit the data depends heavily on the parameter space covered. If I had

explored only a more restricted subset of plaid speeds the angle model might do just as well, or better, than my model for all subjects. My model might instead do significantly better than other models at a broader range of speeds. Similar to what I mentioned with the earlier work on plaid perception, I cannot yet predict with certainty how any model would fair in predicting the data for very different plaids.

I believe that a useful approach for extending these results would be to test specific aspects of my model with psychophysical measurements. I would generate additional model predictions, perhaps for cases where there is existing psychophysical data in the literature. It would also be useful to test specific elements of my model directly, such as by measuring variability in the priors and likelihoods across subjects and stimuli, as I have discussed.

My model does not connect to any specific set of mechanisms, although it is based on known properties of those mechanisms and it suggests the nature of the probabilistic computations that they perform. Linking Bayesian models to physiology models is an important direction. Although this has been discussed recently and some ideas have been described additional work is needed [149].

Another important direction, which has surprisingly not been thoroughly investigated to date, is what role cells with different tuning properties play in forming plaid percepts. Are pattern or component cells neural correlates of coherence or transparency? What is their activity during coherent or transparent percepts? Are pattern direction-selective cells active only during coherent percepts, component direction-selective cells only during transparent percepts? Relatedly, do cells with particular tuning properties, and which are neural correlates of either percept, have different projections to higher areas? Physiological studies could be done

to address these questions. And experiments that involve altering neural activity during plaid perception, as was done with the perception of drifting dots, could be useful.

Chapter 3

Adaptation-induced repulsion is away from pattern directions

Adaptation to moving patterns, biases the perceived DoM of subsequently viewed patterns, away from that of an adaptor. How does this apply for a drifting plaid, which contains motion in more than one direction? I considered these hypotheses: repulsion will be away from the DoM of both components, each of which should activate distinct directional mechanisms; repulsion will be away from the perceived DoM of the adaptor. The "component" hypothesis is motivated by the thought that perhaps it is individual component mechanisms (e.g., either direction-selective V1 cells, or MT component cells) that are being adapted, and that are responsible for subsequent judgments of grating direction. Then you would expect repulsion away from both plaid components. The "perception" hypothesis is motivated by the thought that what you perceive corresponds to the most active units, which are thus the ones that adapt the most. I tested these hypotheses in a series of psychophysical experiments in which subjects indicated the DoM of

drifting square-wave tests after adapting to square-wave plaids and several related stimuli separately. Subjects' perception of the DoM of drifting tests was shifted away from the pattern directions of the plaids, regardless of how they appeared to move. The pattern of perceptual shifts could be explained by an additive mixture of component and pattern adaptation. For some subjects, adapting to transparent random dots caused repulsion away from two different directions. I infer from these results that mechanisms for representing pattern motions adapt regardless of whether they are perceived.

3.1 Introduction

A simple model for visual adaptation is that it results from gain changes in mechanisms that underlie the perception of a stimulus following prolonged exposure to it. Physiologically this can be thought of as a population of neurons which after representing a stimulus for a long time, reduce their responsivity. This reduction is thought to be in proportion to their sensitivity to the stimulus or to particular parameters within it. The reduction shifts the value of the encoded stimulus parameters away from those to which the system had been exposed.

Under these assumptions, repulsive perceptual shifts for a particular stimulus parameter have been taken as evidence of the existence of mechanisms that encode that parameter. Blakemore and colleagues, for example, found that the perceived spatial frequency of a sinusoidal grating is shifted away from the frequency of an adapting grating. They inferred from this that there are visual mechanisms tuned for spatial frequency. Subsequent studies provided further details on these spatial-frequency tuned mechanisms.

I explored the perceptual consequences of adaptation in mechanisms that represent the direction of visual motion. Much is known about these mechanisms. For example, neural responses to plaid stimuli are frequently tuned to motion that is orthogonal to the orientation of the components. A subset of MT neurons, though, is tuned to the 2D pattern motion. Relatedly, plaids can be seen either as transparent components sliding over each other in different directions or as a rigidly moving coherent pattern.

A simple two-stage model arises from these observations in which V1 is thought of as linear and incapable of representing the true motion. MT responses are thought of as informative (nonlinear) and capable of representing the real motions [7, 114, 9, 60, 62]. A method for doing the integration involved in forming a representation of pattern motion is IOC. Each component has a corresponding constraint line in velocity-space. These represent the velocities consistent with the component motions within a finite region of visual space. A theoretical implementation of IOC involves summing the energy passed through a set of spatial temporal filters, each of which represents a portion of a plane in frequency space. Despite significant progress along these lines, I do not yet know how component and pattern direction-selective neurons respond during the perception of a plaid.

I was interested in the perceptual consequences of adaptation to plaids that are perceived to move in these different ways (Figure 3.1). I took advantage of the effects of component and pattern speed on plaid perception in choosing adaptors that appeared to move one way or another for long durations. I used methods advanced by Schrater and Simoncelli, building on earlier work by Sekuler, Blakemore and colleagues, to measure the pattern of shifts in perceived direction of motion that results from adapting to a drifting plaid [66, 65, 68, 69, 83]. I ex-

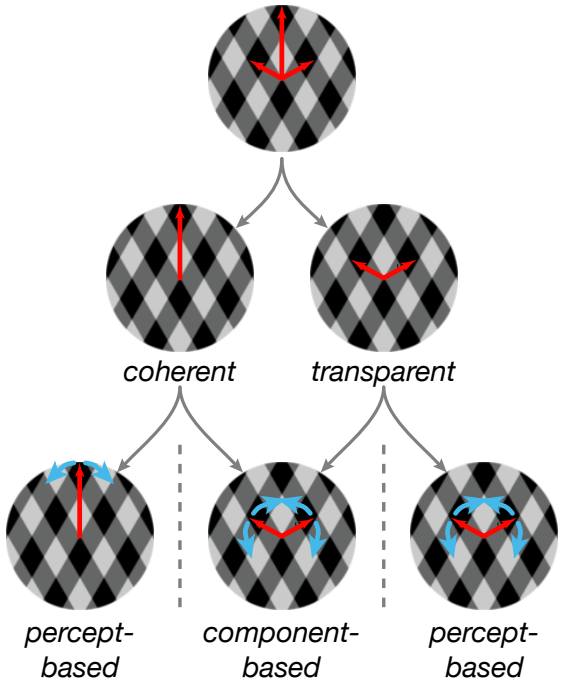

Figure 3.1: An upward moving plaid can be seen as moving coherently or transparently. I considered two hypotheses for the effects of adapting to plaids that are perceived as moving in these ways. One was that the pattern of directional repulsions would follow the percept, such that it is away from the pattern direction when the plaid is perceived to be moving coherently, and away from the component directions when it is perceived to be moving transparently (the outside plots). The second hypothesis was that the directional repulsions would be away from the component directions (the middle plot).

plored adapting not only to coherent and transparent square-wave plaids, but also their components, and to patterns of transparent random dots. My results from these psychophysical experiments, along with a simple least squares analysis, led to a surprising conclusion about the connection between what is perceived and what is adapted, when viewing moving stimuli.

My results are complimentary to previous results on the perceptual consequences of adapting to transparent stimuli [150, 87, 151, 152, 153, 88, 90, 89].

These studies have, in general, considered the problem in terms of the classical movement aftereffect (MAE). A central finding from them is that adapting to a transparent stimulus with motions in different directions causes the perceived direction of the MAE to be unidirectional, and in the direction opposite of the vector average direction of the adaptor [151]. But it is not possible to predict, from this, or other previous results, what the effects of all transparently moving adaptors will be on the perceived direction of drifting tests. Without a clear framework for how adaptation effects with tests with speeds of zero relate to those with positive speeds, predictions of one from the other cannot be made.

3.2 Methods

3.2.1 Apparatus

Stimuli were generated on a dual quad-core 3 GHz Intel Xeon Apple Mac Pro using Expo (version 1.5.1.390) and displayed on a Totoku Procalix 21" color CRT. The monitor refresh rate was 100 Hz and it had a resolution of 1024x768. Its maximum luminance was 80.8 cd/m^2. It was calibrated and gamma-corrected. The subject's viewing distance was fixed at 61 cm and head position was stabilized using a chin rest. At this distance, monitor pixels subtended 0.035 deg.

3.2.2 Stimuli

My main stimuli were upward drifting foveal square-wave plaids. They were 4 deg in diameter with centered circular fixation marks. The spatial frequency of both components was 1.4 c/deg. The overall contrast of the stimulus was 0.8,

with 0.4 for each component. Stimuli were cosine-windowed in space and time. The stimulus would smoothly transition to full contrast within 1% of the time for epoch in which it was presented.

3.2.3 Subjects

Two male and two female subjects (from 26 to 30 years of age) with normal or corrected-to-normal vision participated in my experiments. They were chosen without regard to gender, race, or ethnicity and were compensated for their participation. Three subjects were unaware of the questions being explored in these experiments, but had previously participated in similar psychophysical tasks. One subject was an author. All subjects viewed stimuli binocularly. My experiments were approved by the human subjects committee of New York University. All subjects signed an approved consent form.

3.2.4 Choosing plaid adaptors

The first step in my experiments was to select two plaids for each subject, one that appeared to move coherently and one that moved transparently. Some subjects were not familiar with these percepts and I showed them a brief demonstration of plaids that strongly favored one or the other of them. I did two experiments with a simple forced-choice design in which subjects indicated whether a full contrast upwardly moving plaid looked coherent or transparent. In the task, a plaid appeared for 3 s after which there was a 1 s response period, during which the fixation mark remained and subjects pressed 'C' or 'T', followed by a 2 s blank period.

I chose a slow component speed for which subjects saw plaids to be coherent.

I measured responses for plaids with a fixed plaid angle of 130 deg and with component speeds from 0.5 to 5 deg/s in 0.5 deg/s increments (10 component speeds). I anticipated that in the next series, in which I varied the pattern speed, percepts would transition from coherent to transparent at faster pattern speeds (larger plaid angles). In other words, I attempted to chose a component speed that was highly coherent, but still close to the transition from coherent to transparent along the angle from slow to fast component speeds. This allowed me to choose a transparent plaid when I made measurements along the plaid's pattern speeds. One subject did not see the plaids at the slowest component speeds as being highly coherent, so I chose a slow component speed, assuming that the subject's responses might be coherent at slower pattern speeds. I acted as if their transition along that angle occurred at slower speeds than the ones I explored.

In the second experiment, I chose two different pattern speeds, one for each percept. For the component speed I set in the previous experiment, I measured a subject's percepts at plaid angles from 30 to 150 deg in 30 deg increments (5 plaid angles). Since I measured at fixed angles for all subjects, the pattern speeds they were shown were not the same. The fastest pattern speeds that I showed were for subjects with faster component speeds. As much as was possible, I chose pattern speeds that were close to the transition from coherent percepts at slow speeds to transparent ones at fast speeds. To some extent, I biased the pattern speed for the coherent adaptor to be slower than it could have been. I wanted to ensure that the plaid would appear strongly coherent for the full duration of my presentations of it. In some cases, I repeated these measurements on different days with the same subject. This allowed me to confirm that plaids with the parameters I had chosen still evoked the intended percepts.

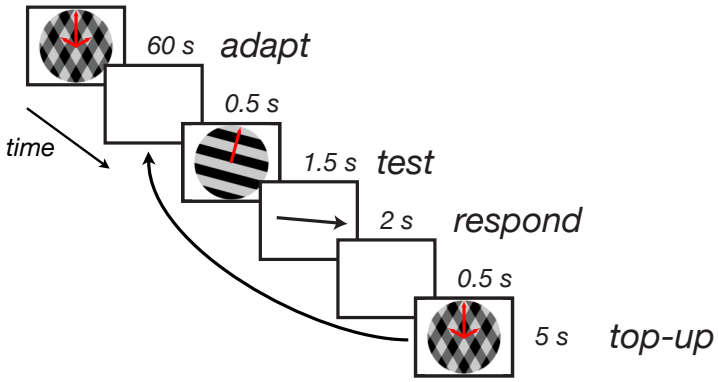

Figure 3.2: A diagram for my method of adjustment task for measuring the directional repulsions. Time flows from the upper left to the lower right. Subjects viewed an upward drifting square-wave plaid for 60 s. After a brief blank, they viewed a square-wave test grating moving in some direction around the circle. They indicated the direction they perceived the test grating to be moving by adjusting the orientation of an arrow. Shortly after that the adaptor was repeated for 5 s and the sequence repeated.

3.2.5 Measuring directional repulsion

I used a method of adjustment task to measure the directional repulsion that results from prolonged viewing of plaids and other stimuli (Figure 3.2). The task began with an initial adaptation period for which the adaptor stimulus (i.e. an upwardly moving plaid) appeared for 60 s. After 0.5 s of blank, a drifting square-wave test grating appeared for 1.5 s. Immediately thereafter an adjustable arrow appeared for 2 s, which the subjects rotated to indicate the direction that they perceived the test grating to have moved. The starting orientation of the arrow was randomized from trial to trial. They were instructed to indicate the direction of the motion in particular and not, for example, the direction orthogonal to what they perceived the orientation of the test to be. The spatial windowing on the drifting test may have minimized the possibility that they could have simply indicated the point at the edge of the aperture where the gratings converged. They could

adjust the pointer until the end of the 2 s, the time from which I recorded their response. This response period was followed by a 0.5 s blank and then a 5 s 'topup' repeat of the adaptor. This was a full trial. The sequence would repeat thereafter, starting with the 0.5 s blank. A small circular fixation mark was presented during all periods of the adaptor, the test grating and the response. It was not presented during the blank periods.

Psychophysical responses from all subjects were collected for eight different adaptors: the coherent and transparent plaids; the components for those plaids (four square-wave gratings); upward drifting coherent random dots; and transparent random dots. For all of these adaptors, I used drifting square-wave gratings as test stimuli. The coherent and transparent plaids differed only in their plaid angles. The speeds of their components were the same. The pattern speeds were different, however, and were faster for the transparent plaid. I presented test gratings with only two different speeds, one was equal to the component speed and the other was equal to the pattern speed. For the coherent plaid and its component adaptors, the pattern speed that I matched in the tests was that of the coherent plaid. For the transparent plaid, its components, and all random dot adaptors, the pattern speed that was matched for the tests was that of the transparent plaid. The directions of the tests spanned the circle in 15 deg increments (24 different directions). Also, the dot adaptors that I presented had speeds slightly less than those of the plaids. They were faster for those subjects with faster plaid adaptors.

3.2.6 Analysis

To quantify the extent of the perceptual shifts in direction, I computed the difference between the mean resultant of the reported perceived directions across

trials, with the actual test direction. This gave a single value of the shift for each test direction. I also computed confidence intervals for these shifts by bootstrapping the difference between the mean resultant and the actual test DoMs. I have not shown these intervals in my circle-shaped plots, although I used them in assessing the data. I mostly considered the direction of the repulsive centers, the directions for which repulsive shifts switched from clockwise to counterclockwise or vice versa, in evaluating my hypotheses. For the coherent plaid, for example, if there was only a repulsive loci away from the upward direction, this supported my hypothesis that the adaptation was consistent with the percept, and not with the directions of the component motions.

I also evaluated how well the pattern of perceptual shifts could be explained as an additive superposition of component-based and pattern-based shifts. The predictors for this analysis were derived from psychophysical data for each subject. Component predictions for the coherent plaid and transparent plaid adaptors were determined by adding the shifts for the components for those plaids. This loosely resulted in what one might expect, repulsive shifts away from both component directions. The pattern prediction that I used was from the upward drifting coherent random dots, which likewise caused repulsion away from the upward direction, the extent of which varied across subjects. I used least squares to separately determine the weights for each prediction that, when summed, best explained the plaid data. I used a standard calculation of the coefficient of determination to evaluate the goodness of fit of these models. I have shown the best predictions of the data from these models by simply multiplying the weights that I determined by the component and pattern predictions.

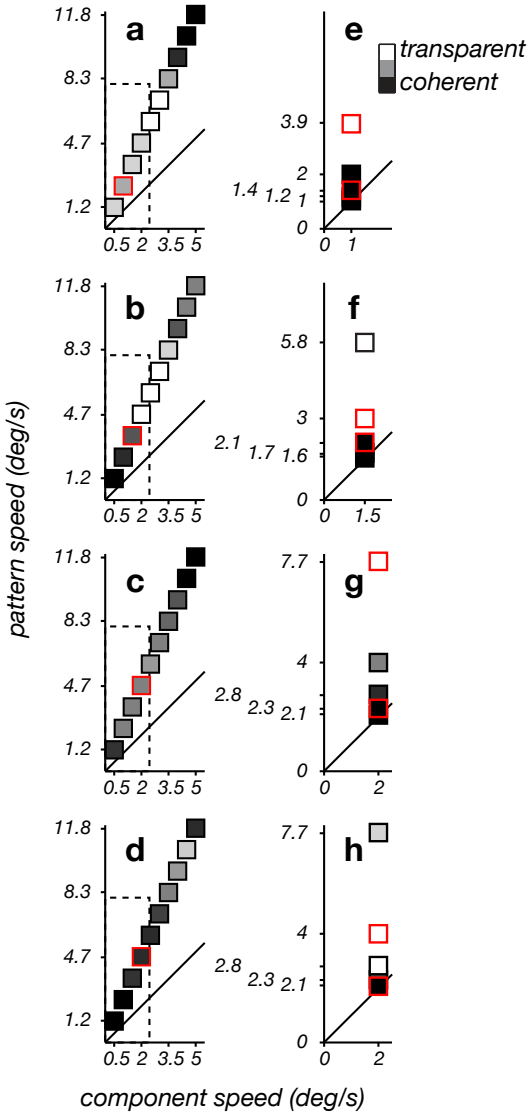

Figure 3.3: Percepts from two different forced-choice tasks, from which I chose the coherent and transparent plaids. Each square represents the ratio of percepts for a plaid with a given component and pattern speed. (a-d) Plaids with the same angle between the direction of their component motions were presented at 10 different component speeds. The square with the red highlight is for the component speed that I selected for the the subsequent plaids. The dashed rectangle is the region of speeds represented in the plots e-h. (e-h) Plaids with the same component speed and with five different plaid angles are shown. The squares with red highlights are for the plaid adaptors.

3.3 Results

3.3.1 Coherent and transparent plaids

Subjects' percepts varied with component speed (Figure 3.3a-d). The whiteness of the squares in these plots corresponds to the ratio of percepts across trials, with white squares corresponding to all transparent, black to all coherent. Two subjects saw plaids at nearly all component speeds as coherent. One subject saw some as transparent for the middle of the range of speeds. The other remaining subject saw those plaids with slow component speeds as coherent. As mentioned previously, these data were quite variable across subjects, and it was not possible to follow a strict rule in choosing the component speed for subjects' plaids. As much as possible, I tried to chose one that was in a region of some transparency or just slower. The red highlighted squares in these plots are for the component speed that I selected for plaid adaptors. When all responses were coherent, I chose a component speed that was slightly faster than what was chosen for other subjects.

In the next experiment, which I collected immediately after the previous one, I varied the pattern speed (Figure 3.3e-h). Transparency always increased with faster pattern speeds. The two red highlighted squares in these plots are for the two pattern speeds that I used for the plaid adaptors. In all cases, I was able to choose a plaid that was strongly coherent and another that was strongly transparent. Curiously these data show inconsistencies in the percepts for matched speeds with the previous component speed series. For example, subject 4 saw plaids as transparent at pattern speeds in the second series, which were seen as coherent in the previous one. I do not know the origin of this variability.

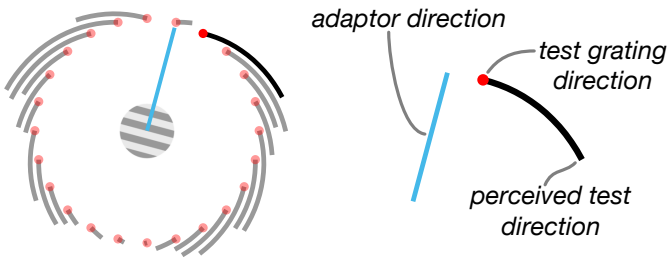

Figure 3.4: Results of my method of adjustment task in a circle-shaped plot. Some elements of the plot are shown to the right. The blue line indicates the DoM of the adaptor. The red dot is the actual DoM of a drifting test grating. The arced black segment shows the extent of the difference between the actual direction of the test and the perceived direction. That these segments extend out is for illustration purposes and does not, for example, indicate any change in the perceived speed of the tests.

3.3.2 Component and pattern predictions

Figure 3.4 shows some of the results of my method of adjustment experiments in a circle-shaped plot. The red dots show the direction of the moving square-wave tests. The arced black segments show the path from the actual test DoM to the perceived one (which is the mean resultant of the pointer directions). The blue line segments at the center of these plots indicate the directions of the adaptor. Where there is a single direction, such as for the upward drifting dots, there is one segment. Where there is more than one direction, such as for the coherent plaid, all directions are shown. For that plaid, the directions of the component and the pattern velocities are shown. I use solid segments for elements of the adaptor that are perceived and dashed ones for those that are not. For the least squares fits, I show all three segments as solid, since they are formed from the component and pattern predictions.

I formed predictions for the component adaptation and the pattern adaptation by measuring the directional biases with different adaptors (Figure 3.5a-h). I did

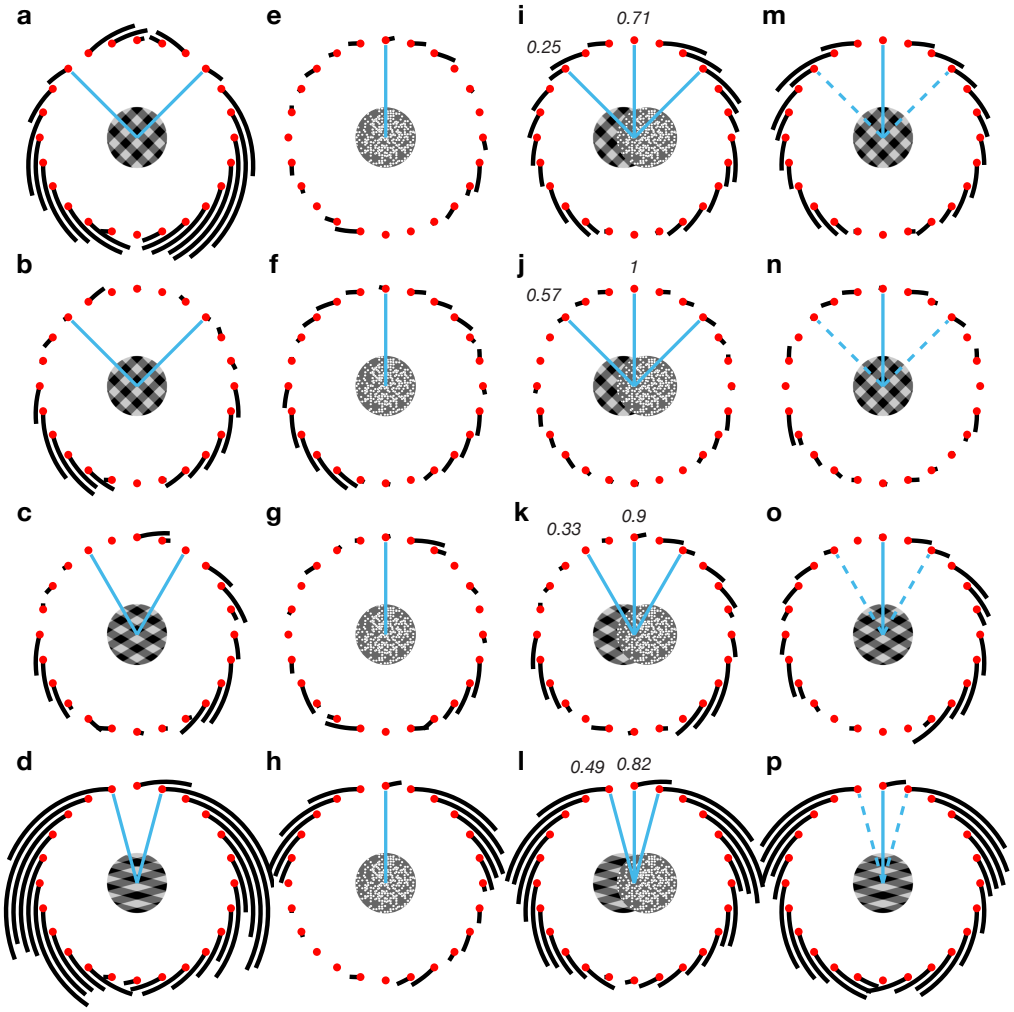

Figure 3.5: (a-d) The component predictions from the sum of adapting to different component motions. (e-h) The pattern predictions from adapting to upward drifting random dots. (i-l) The superposition of the two predictions that best matches the real data for adapting to a coherent plaid. The numbers are the weights of the component and pattern predictions from that superposition. (m-p) Adapting to a coherent plaid caused repulsion away from the pattern direction.

not know how the plaid adaptation data would relate to these predictions. The predictions provided a baseline of what the underlying adaptations might be like. There are different component predictions for the coherent and transparent plaids. For all subjects, I measured the perceived directions for each component by itself as an adaptor. I summed the responses for the leftward and rightward components.

The component prediction for the coherent plaid generally has repulsion away from the DoM of both components (Figure 3.5a-d). The exception for this was subject 4, where the plaid angle was only 30 deg. For that subject, there was repulsion away from the upward direction. There is variation in the scale of the shifts across subjects, with some being much larger than for the others. The shifts are often quite large and sometimes cross over oblique directions. There are also some cases in which the shifts are insignificant. They are generally symmetric, although they are sometimes larger for clockwise or counterclockwise shifts for a subset of the tests. The coherent plaid component prediction for subject 1, for example, has larger shifts for rightward tests, which are shifted in a clockwise direction.

The component prediction for the transparent plaid also has repulsion away from both component motions (Figure 3.6a-d). These repulsions are generally more significant than for the component prediction for the coherent plaid. There are large shifts away from both component directions, which are quite spread apart. For three of the subjects, the DoM of the components were only 15 deg above straight left and right. There are also larger downward shifts for the test directions in the lower quadrants for two of the four subjects as compared to the coherent adaptor. The first subject seems to have slightly smaller downward shifts.

I used upwardly moving coherent random dots for the pattern prediction. I

used the same pattern prediction for the coherent and transparent plaids. They caused repulsion away from the upward direction for all subjects (Figure 3.5e-h). This repulsion is comparatively not very large, but they are always away from the upward direction. Most of the repulsion from this adaptor is confined to upwardly moving tests. Subject 4, for example, has minimal shifts for downwardly moving tests. Subject 2, though, has clear attraction towards the downward direction.

3.3.3 Coherent plaid adaptation

Adapting to coherent plaids caused repulsion away from the upward direction (Figure 3.5m-p). The locus of the adaptation is the same as the perceived pattern direction of the plaid. None of the subjects has clear repulsion away from the component directions. Subjects' percepts are shifted to a greater extent for upward directions, especially those that are close to the upward direction. This is clear with Subject 4 where the three closest directions to the upward direction are shifted nearly to straight left and right. The downward drifting tests are also shifted. They appeared to be moving in a downward direction, although not quite as much. Most subjects had biased percepts of DoM for all directions. This differs from their pattern prediction with upward drifting dots, which has less of an effect for downward tests. There are some tests which are inconsistent with this. Subject 2, for example, has upward shifts for two tests, one leftward and one rightward. These shifts do not conform to either prediction. They might result from an attraction that is reversed from the upward repulsion.

The combination of the predictions that best matches the data are very similar to them (Figure 3.5i-l). Like the actual data, there is repulsion away from the upward direction. This results from a significant weighting of the pattern predic-

tion. The downward shifts might result from some weighting of the component prediction. The shifts are larger for the rightward downward shifts for subject 3 in both the psychophysical data and the predicted data.

For this adaptor, the shifts are away from the perceived direction of the stimulus, which supports the hypothesis that biases from adaptation follow the percept. I did not see evidence for the component-based hypothesis, although the attractive shifts towards the downward direction may be consistent with it. An alternative possibility, is that coherent plaids affect more directions than my prediction for the upward pattern adaptation. They may affect the opposite direction differently than my prediction. They might generate greater disinhibition for opposite directions than the upwardly moving dots. My modeling results, which I turn to in a subsequent section, suggest that a mixture of the component and pattern predictions can account for these results.

3.3.4 Transparent plaid adaptation

Surprisingly, transparent plaids caused repulsion away from the upward pattern direction (Figure 3.6m-p). The data for all subjects includes repulsion away from the upward direction even when the component directions are quite far apart. This means that there is adaptation for the pattern direction even when it is not perceived. This result is inconsistent with both of my hypotheses. There is little or no repulsion away from the component directions. There are, however, a few test directions which do not show shifts away from the upward direction. It is not clear whether these could be the result of component-based repulsion.

The pattern of DoM biases is similar to those for the coherent plaid adaptor, even though the DoM of the components are different. There are shifts in the

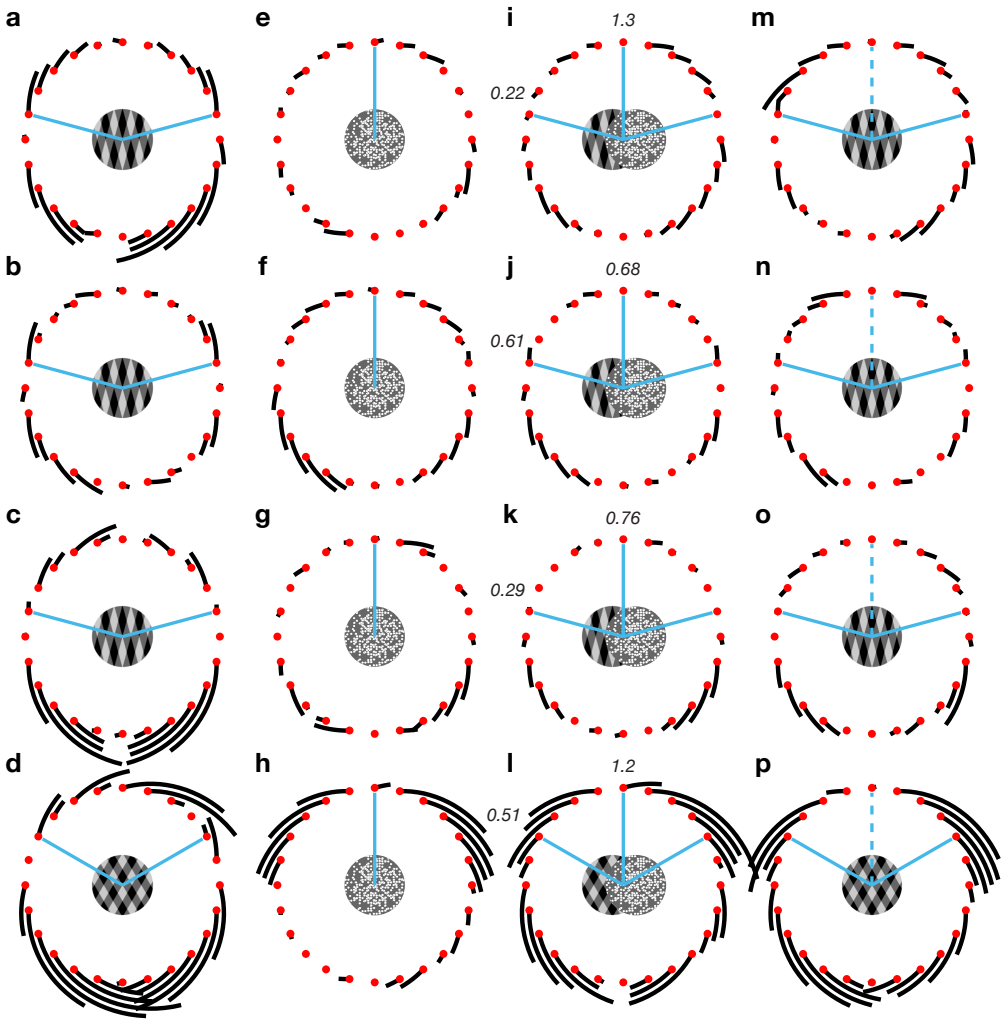

Figure 3.6: (a-d) The component predictions from the sum of adapting to different component motions. (e-h) The pattern predictions from adapting to upward drifting random dots. (i-l) The superposition of the two predictions that best matches the real data for adapting to a transparent plaid. The numbers are the weights of the component and pattern predictions from that superposition. (m-p) Adapting to a transparent plaid caused repulsion away from the pattern direction.

perceived direction for tests in nearly all directions, including those in the lower quadrants. There are larger shifts for tests that have some upward directional component, and which are proximal to the repulsive center. The shifts are smaller for tests with a downward component. Whatever the origin of this attraction to downward directions, it is not significantly greater for the transparent plaid.

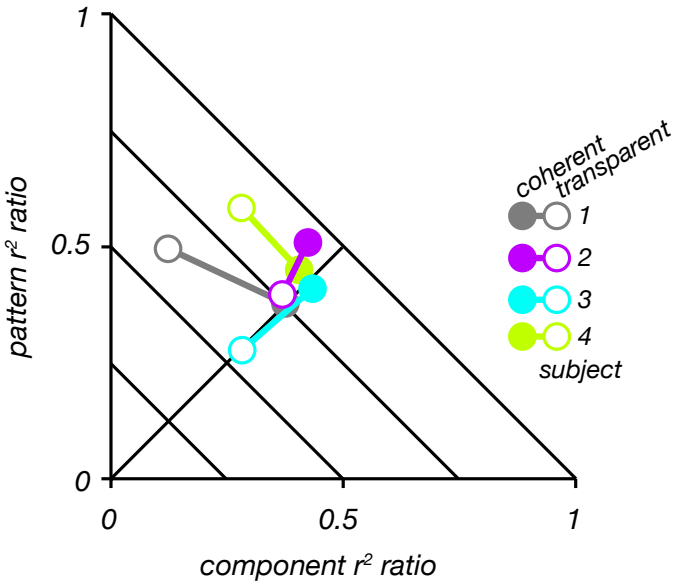

Figure 3.7: The normalized coefficient of determination of the component and pattern predictions for coherent and transparent plaid adaptors. Solid dots are for the coherent plaid; open dots for the transparent plaid. For all subjects, the component and pattern predictions account for roughly equal amounts of variance. For two subjects, the predictions explain less variance for the transparent plaid. For the other two subjects, the pattern prediction accounts for even more variance than in the coherent case.

For two of the subjects, the least squares fits show that the weight of the pattern prediction is greater, and the component prediction is less, for the transparent case (Figure 3.7). For the other two, the relative weights of the two predictions remain about the same, although the goodness of the fit decreases (they are closer to the origin). The component-based predictions do not predict the data well for

either plaid adaptor. They wrongly contain repulsive shifts towards the upward direction, not away from it. They are especially poor predictions in the transparent case since the component DoMs are further spread apart. They have even larger shifts towards the upward direction for more tests. I tried a different transparent adaptor to evaluate whether there is a limitation on the number of repulsive loci. This could be a constraint of the system for there to be a single such locus of repulsion. This constraint could account for the repulsion for the transparent plaid.

3.3.5 Transparent dots adaptation

I collected a set of data for an adaptor of transparently overlaid drifting random dots. The DoMs of the dots were the same as for the component motions of the transparent plaid. Two of the subjects had repulsive biases away from both directions of random dots (Figure 3.8). In these subjects, there is, for example, perceptual biases towards the upward direction. In one other subject, there is no clear pattern of biases. There may be some repulsion away from the directions of both sets of dots as well as some repulsion away from the upward direction. Results for this subject seem generally noisy and there are several tests for which there was no effect. The last subject had repulsion away from the upward direction. The shifts for this subject were smaller than for the upwardly moving plaids. Also, these shifts did not show clear attraction toward downward directions, and may actually have some repulsion away from them.

The results from the two subjects that showed similar repulsive shifts away from both transparent dot directions are consistent with my previous results, in that they are consistent with adaptation in pattern-tuned mechanisms. In this

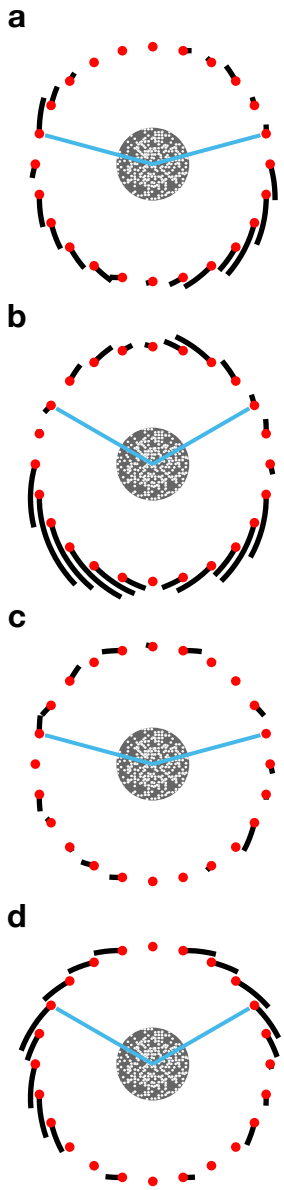

Figure 3.8: Adaptation-induced biases from transparent random dots. (a-b) In two subjects, adapting to transparent random dots causes repulsion away from the directions of both sets of dots. (c-d) In the other subjects, there is less repulsion away from those directions.

case, the percept of the adaptor that is likely to be formed by those pattern-tuned mechanisms is the same as what is adapted. The results from the forth subject are inconsistent with those for the other subjects. This subject may have adaptation for the component motions that is masked by stronger adaptation in mechanisms representing the upward direction. Their repulsion away from the upward direction could also be the result of both components having significant upward components of motion. I could have acquired additional confirmation that this subject saw these dots to be moving transparently, although I believed they did and they did not indicate otherwise. This result confirms that there is not a limit on the number of repulsive loci in general. It provides additional support that adaptation to moving stimuli causes adaptation in pattern-direction tuned mechanisms. The repulsion coincides with the perceived directions when those directions are the same as the pattern directions.

3.4 Discussion

My results suggest that mechanisms for representing pattern motion adapt, and affect subsequent perception, regardless of the way in which adapting stimuli are perceived. I have shown that: (1) adaptation to plaids that are perceived to move coherently or transparently cause repulsion away from their pattern directions; (2) adaptation to these stimuli is inconsistent with hypotheses that the pattern of perceptual shifts is always away from the component directions of motion or away from the perceived DoM; (3) the overall pattern of shifts can be explained by an additive mixture of component- and pattern-based adaptation; (4) the ratio of the weights for the pattern prediction and the component prediction that best explain

the data are either approximately equal, or increase, which is to say they become more pattern-like, for transparent plaids relative to coherent ones; (5) adaptation to transparent random dots causes repulsion in perceived direction away from the directions of the two populations of dots for some subjects. It is possible that there are more subtle differences between adaptation to coherent and transparent plaids, such as the amplitude of the shifts or the extent of the effects in different directions, although they did not stand out in my data.

Neurons that represent pattern motion might adapt in ways that component direction-selective ones do not. The same can be said of neurons that decode representations of those motions, which might adapt for representations of pattern motions, but not component motions. I know of no evidence that supports either of these possibilities and question how such differential adaptability could arise.

I believe that the specificity of stimulation evoked by a visual pattern partly influences the magnitude of the adaptation it causes. Following that, the repulsion away from pattern directions could be explained by there being less uncertainty about it. But component direction-selective mechanisms should be activated with a specificity similar to the pattern direction-selective ones. And variability of measurements for the component motions is likely to be less than that for pattern directions. They may inherit uncertainty for both components, although the representation of the pattern motion may also involve overlapping representations formed by a broader population of cells. In the case of the transparent plaid adaptor especially, there is no reason to believe that the pattern direction, which is unseen, would evoke a more specific representation.

The connection between the mechanisms for representing visual motions and perceiving them is relevant to these results. If, for example, only pattern direction-

selective mechanisms connect to, or are a component of, networks for whom neural activity corresponds to perception, adaptation in them only, would give rise to changes in perceived direction and speed. In other words, mechanisms which represent component motions might be adapting, but not transmitting their modified responses to mechanisms that form the percept of the DoM of the tests.

Electrophysiological or neuroimaging studies could help to disentangle these possibilities. An obvious, but challenging, first step might be to work towards a better model for neurons in MT, and other relevant visual areas, that respond during the perception of plaids, gratings and dots. It could be that only the responses of pattern direction-selective neurons correlate with the percept. Studies in which those responses are manipulated, as was done in other studies of MT, might be able to establish a causal link between neurons with certain sorts of tuning and the perception of visual motion.

In interpreting my results, I have assumed a simple model of motion adaptation, in which mechanisms selective for something have the gain of their responses reduced following prolonged exposure to it. This assumption was applied in earlier studies, such as those which showed repulsion in the perception of spatial frequency and direction, and inferred that there exists channels selective for those properties. In my case, I knew that there were mechanisms tuned for component and pattern motion and wanted to know the reverse, which was whether the repulsion would be consistent with one or the other of them.

That said, the means by which adaptation occurs, is likely to be significantly more complicated and may not be limited to gain changes. Recent studies on adaptation in MT suggest that neurons whose tuning is centered on an adaptor, narrow their tuning at the point of the adaptor [154]. The same occurs for neurons

with tuning flanking the adaptor. These neurons also have attractive shifts in their tuning towards the adaptor, so the centers of their tuning curves change [154]. Simulations of these neurons suggested that these changes, not gain changes, are capable of generating the associated perceptual effects.

It is likely that preferred and flank adaptation of component- and pattern-tuned mechanisms plays a role in the perceptual effects that I measured with drifting plaids. It is not yet possible, though, to say which sorts of adaptation are contributing. For sine-wave gratings there are significant flank adaptation effects for directions further than 90 deg away from the preferred direction of an MT cell [154]. To constrain these possibilities, it should be useful to measure the responses of MT cells after adapting to drifting plaids. What is the shift in the preferred directions of pattern and component cells after adapting to plaids moving in the preferred, flank, end-of-flank, and null directions? How is the bandwidth of the tuning of the cells affected by adaptation in those relative directions?

The levels at which adaptation occurs within the visual system has only recently been considered. How do changes in one population of neurons propagate to others in downstream areas? For the coherent plaid adaptors that I presented, for example, there could be adaptation in mechanisms tuned to a variety of different properties of the stimulus, such as its contrast, orientation, size or color. The task that subjects are asked to do, and the test stimuli used, should constrain the relevance of some stimulus properties, and therein minimize the consequences of adaptation from them. But if several properties are represented simultaneously they may be adapting in parallel. What are the interaction effects of parallel adaptation in different channels?

The motion aftereffect (MAE) has been shown to depend on the type of test

stimuli used to measure it. The classical MAE, in which a stationary test is seen to move as a result of motion adaptation in the opposite direction, is different than the effect that is observed with flickering or moving test stimuli. The MAE with stationary tests is spatial- and temporal-frequency tuned, whereas it is mostly independent of spatial-frequency, and is tuned for speed, with flickering or moving tests. These differences have been taken as evidence that the two types of tests are probing the adaptation of different mechanisms.

I have not focused on a singular consequence of adaptation, such as inducing motion in stationary tests, which is the norm in studies of the MAE. I was interested in adaptation as a way to uncover representations of the perception of different sorts of visual motion. Since adaptation frequently causes perceptual changes throughout the 2D velocity space that the visual system represents, there seems to be little reason to consider the effects only on tests with zero speed. The effects may be of greater variety or consequence elsewhere in the space.

Previous studies have found that transparently moving adaptors cause singular aftereffects for stationary or flickering tests. These shifts might be singular in their DoM, for example away from the vector average of the adaptor. This could result from repulsion in velocity space away from that direction (along the path from the pattern direction through the origin).

For the stimuli I explored, uncovering repulsion in more than one direction may be less likely with stationary or flickering tests. It is unclear how multiple repulsions could be seen with them. I do not know if it is possible for them to appear to simultaneously move in different directions. They might move in one direction on some trials and a different one on others, switching back and forth between two different directions. Since my measurements were for tests moving

in many different directions, at speeds related to those of the adaptors, multiple repulsive loci should have caused clear patterns of perceptual biases if they are generated. My results with transparently moving random dots demonstrate that adaptation can cause multiple repulsive loci.

Part of my motivation in using drifting square-wave tests was that I believed that if component-based adaptation results from plaid adaptors I was most likely to see it with tests that are similar to those components. It actually seems like the type of moving tests used may not matter very much, and I believe that my results would be similar with some other types of tests. Adaptation causes repulsion in coherent plaids–they are shifted away along a vector eminating from the adaptor's position, through the test's pattern direction. I do not know what the effects would be with transparent plaid tests or other transparently moving tests, although it would likely result in repulsion away from their pattern directions. I have speculated about several issues related to these results and I conclude by suggesting some specific studies that could follow them. I also highlight some tests of my interpretation.

An important goal is to develop a fuller and more general normative model for adaptation from moving stimuli. These psychophysical results, and others, should be incorporated into Bayesian models for motion perception. Current work on adaptation in this context has shown that adaptation causes changes in the likelihood function, which is an internal inference of a parameter or a stimulus given some sensory input. As might be expected, adaptation causes increases in the signal-to-noise ratio of measurements by adjusting the operational range of the measurement stage so that it conforms to the input range of the stimuli. It might be possible to add to this sort of model by working backwards from the pattern

of perceptual effects that I observed to changes in tuning for localized regions of velocity space.

More precise measurements of changes in speed and direction for drifting gratings, plaids and dots should be made. These could address whether adaptation to gratings, for example, causes circular repulsion in adjacent velocities along the gratings constraint line. Transparent dots should cause repulsion away from both dot velocities, so that gratings near those loci with faster speeds appear to move at higher speeds and vice versa. The perceptual effects for directions opposite the pattern direction of the adaptor are also of interest. Disinhibition likely contributes, as the amount of inhibition is decreased by a reduction in the responses for the opposite direction.

Finally, an important test of my interpretation, that pattern direction-selective mechanisms adapt regardless of what is perceived, could be made with adaptors that are designed to reverse the effects of the initial adaptation. Stimuli in which energy is subtracted in specific locations from a block with energy in many directions at many different frequencies could be used to nullify the effects of different adaptors. The effects should depend on the strength of the reverse-adaptors. An important prediction from my results is that these hole-adaptors would negate the effects of plaid and random dot adaptors when subsequently presented adaptors lack energy in the pattern-directions. They should have minimal effects when their holes are centered on other locations, including those for the component motions.

Chapter 4

Conclusions and future work

To this point, I have described my investigations of three issues related to the integration of visual motion. In the chapters associated with each of these, I have included summaries of my results as well as a discussion of what they mean and how they might be improved upon. I will conclude by briefly restating my findings, and by suggesting a few proximal directions that could be pursued in future work. To me, the latter is particularly important. My research has provided only partial answers to the questions I set out to explore–though it also creates more questions.

4.1 The scale of visual motion

My first set of results showed that MT neurons do not represent long-range apparent motion, which is to say that the neurons' responses are unaffected by such motion. This is consistent with their retaining the scale of the directionality found in their upstream inputs. Thus, when a person's percept of a stimulus is of widely spaced motion, MT's responses will be uncorrelated with their sense of motion. This may seem puzzling given the view that responses in MT underlie a sense of

motion. This begs the question why MT neurons are computing something that is unnoticed, or is disconnected, from any obvious function? Also, if MT neurons are not representing long-range motion, what I have called global motion, which neurons are?

To find out, one might record from many other visual areas, including areas in the ventral stream, with local-global motion stimuli. It would be advantageous to have some idea of where to look. And the number of possible areas to consider could be constrained in part by the general characteristics of the neurons in the area, such as their RF size. It is possible that neurons which represent global motions have little or no directionality for local motions, and therefore might be thought of as not playing a role in representing motions.

Another approach for identifying areas that represent global motions might involve measuring the effects of adapting to local-global stimuli with fMRI. Stimuli could be used in which the direction of either the local or the global motion is randomized. Adaptation effects from different combinations of directed, random or static, local and global motion could be used to isolate regions that are selective for either scale of motion. A series of nested differencing metrics, based on changes in responses, could be used. These would be similar to the ones I used to quantify the scale of directionality in neurons. MT+, and other dorsal stream areas, might be selectively adapted only by local motions.

If there are individual neurons that are representing global motion, how are they doing so? One possibility is that insensitivity to local motion is made by averaging the responses of sets of many different directionally-tuned neurons. These could overlap in the portions of visual space which they represent. From the output of these, could be added enhancement for a particular sequence of excitations, and an

inhibition for the opposite sequence. This would form large, oriented, space-time filters, which would be sensitive to the detailed, oriented content of global motions in the frequency domain.

Representing local and global motions is a subset of the general issue how humans see motions at different scales. Although there has been some progress on what those scales are and how representations of them are formed, more work is needed. Remaining questions include, what are the boundaries of those scales and what establishes the boundaries? What is the origin of representing different scales of motion, which is related to the functions they subserve? Relatedly, what are the correlates of different scales of motion in natural scenes, or in the experience of everyday life? It will be important to know more about how motions at different scales interact in the images themselves, and how perceptions of them are interrelated.

4.2 To integrate or to segment

My second set of results showed that the perception of drifting plaids is influenced by their component and pattern speeds. In addition, a Bayesian observer model that includes two priors, which impart preferences to see slow and singular motions, can account for the relationship between a plaid's speeds and whether its motions are integrated or segmented.

Improvements in my model might include an expanded, and more realistic, front-end, as I have described. By front-end, I mean the transformations involved in moving from the stimulus to the likelihoods. Also, a better estimate of both priors, the one for velocity and for singular precepts, is needed. Better estimates would

mean fewer assumptions have to be made. Experiments done with estimating the speeds of 1D motions provide a defined method for how to estimate the priors and likelihoods throughout velocity space. These could be extended to measure the 2D velocity prior. It would also be useful to have estimates for that prior with other stimuli, such as square-wave gratings, and at different eccentricities. Additional questions include whether these priors extend to other tasks. Is the prior to see singular motions, for example, based on a preference to see the fewest motions in all possible cases, or is it a preference only to see one instead of two motions?

These questions have to do with the sorts of computations involved in combining different motions, and more work could be done on how those computations are implemented. There is some evidence that the processing that corresponds to forming likelihoods of motions can done by neurons in MT, that it follows the standard motion energy model of MT. But what about the other steps in the model I described? For example, how are likelihood representations combined with a prior for velocity? Is that, or could it be, done in separate pathways? How are probabilistic estimates integrated and compared in neural terms?

It has been suggested that Bayesian priors are related to the sum of the tuning curves among a population of neurons. These tuning curves are thought to relate to the distribution of properties in the natural environment. These ideas need to be tested further. The relation to tuning curves could be explored in simple model organisms, perhaps by studying a simple behavior and manipulating the responses of cells with optogenetic techniques.

A final direction, somewhat distinct from those utilizing Bayesian models, is to establish the relationship between pattern-tuning and the perception of integrated motion. The large body of work that established the link between MT and motion

perception can be looked to for possible experiments, which will help to clarify how motion integration or segmentation is implemented. Are cells with a particular pattern of tuning for component or pattern motions, active or inactive, during one or the other percepts? Could the probability of the coherent percept be increased by increasing the responses of pattern cells? What is the role of unclassified cells? What are the perceptual consequences of inactivating cells with different tuning properties? Could the responses of some subset of cells, with a particular pattern of tuning, be used to form accurate judgements of motion, assuming the organism had access to them?

4.3 Adapting to visual motion

My third set of results showed that adapting to coherent and transparent plaids causes repulsion in the perceived directions of drifting tests away from the plaids' pattern directions. This suggests that pattern-tuned mechanisms are responding and adapting to a visual stimulus regardless of what the percept is.

This finding is puzzling and there are a number of follow up studies that could be done which might provide some clarification. These include the studies I described that relate to the physiological mechanisms for motion representation with different motion percepts. But, more to the point, what underlies the specific adaptation in pattern-tuned mechanisms? As I discussed, it is unclear whether this has to do with pattern cells having different dynamics with respect to prolonged exposure to a stimulus. It is also possible that only cells with pattern tuning impact subsequent perceptions of motion, but why would that be? To address this, I have suggested presenting stimuli that would reverse the adaptation effects of a drifting

plaid. For example, a stimulus in which energy in the pattern direction of a plaid has been removed could be presented after adapting to a plaid. This could be done with both coherent and transparent plaids.

It would also be useful to have better estimates of the full pattern of adaptation effects for plaids, and other standard visual stimuli. I mean, for example, to use forced-choice experiments to measure speed and direction biases throughout velocity space, which are a consequence of adaptation. There are some interesting possibilities that may or may not emerge from such experiments, such as whether there are repulsions away from velocities along the velocity constraint line of a drifting grating. There may also be interesting things occurring in the opposite direction of a moving stimulus, where disinhibition is likely having an effect.

Temporal dynamics are generally omitted from models of motion processing, mostly for lack of systematic measurements and principles. One thought on adaptation is that it relates to the intensity of the representation of the stimulus. There is also a more recent view that it is related to the specificity of excitation elicited by a stimulus for a sub-population of cells. These principles, although quite general, should be developed and tested further. As they are coupled with additional sets of measurements, fuller models of how visual motions are represented and perceived will be developed.

Bibliography

[1] A. Wohlgemuth. *On the after-effect of seen movement*. University Press, 1911.

[2] H. Wallach. Über visuell wahrgenommene Bewegungsrichtung. *Psychological Research*, 20(1):325–380, 1935.

[3] C. L. Fennema and W. B. Thompson. Velocity determination in scenes containing several moving objects. *Computer Graphics and Image Processing*, 9(4):301–315, April 1979.

[4] D. Marr and S. Ullman. Directional selectivity and its use in early visual processing. *Proc R Soc Lond B Biol Sci*, 211(1183):151–180, Mar 1981.

[5] N. Rubin, S. Hochstein, and S. Solomon. Restricted ability to recover three-dimensional global motion from one-dimensional motion signals: psychophysical observations. *Vision Res*, 35(4):463–476, Feb 1995.

[6] N. Rubin, S. Solomon, and S. Hochstein. Restricted ability to recover three-dimensional global motion from one-dimensional local signals: theoretical observations. *Vision Res*, 35(4):569–578, Feb 1995.

[7] E. H. Adelson and J. A. Movshon. Phenomenal coherence of moving visual patterns. *Nature*, 300(5892):523–525, Dec 1982.

[8] B. K. P. Horn and B. G. Schunck. Determining optical flow. *Artificial Intelligence*, 17(1-3):185–203, August 1981.

[9] J. A. Movshon, E. H. Adelson, M. S. Gizzi, W. T. Newsome, C. Chagas, R. Gattass, and C. Gross. The analysis of moving visual patterns. *Vatican City, Italy: Pontifical Academy of Science*, pages 117–151, 1985.

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

[11] E. Mingolla, J. T. Todd, and J. F. Norman. The perception of globally coherent motion. *Vision Res*, 32(6):1015–1031, Jun 1992.

[12] L. Welch and S. F. Bowne. Coherence determines speed discrimination. *Perception*, 19(4):425–435, 1990.

[13] A. T. Smith. Coherence of plaids comprising components of disparate spatial frequencies. *Vision Res*, 32(2):393–397, Feb 1992.

[14] J. Kim and H. R. Wilson. Dependence of plaid motion coherence on component grating directions. *Vision Res*, 33(17):2479–2489, Dec 1993.

[15] G. R. Stoner, T. D. Albright, and V. S. Ramachandran. Transparency and coherence in human motion perception. *Nature*, 344(6262):153–155, Mar 1990.

[16] G. R. Stoner and T. D. Albright. Neural correlates of perceptual motion coherence. *Nature*, 358(6385):412–414, Jul 1992.

[17] G. R. Stoner and T. D. Albright. Motion coherency rules are form-cue invariant. *Vision Res*, 32(3):465–475, Mar 1992.

[18] G. R. Stoner and T. D. Albright. The interpretation of visual motion: evidence for surface segmentation mechanisms. *Vision Res*, 36(9):1291–1310, May 1996.

[19] G. R. Stoner and T. D. Albright. Luminance contrast affects motion coherency in plaid patterns by acting as a depth-from-occlusion cue. *Vision Res*, 38(3):387–401, Feb 1998.

[20] H. Farid and E. P. Simoncelli. The perception of transparency in moving square-wave plaids. In *Investigative Opthalmology and Visual Science Supplement (ARVO)*, volume 35, page 1271, May 1994.

[21] H. Farid, E. P. Simoncelli, M. J. Bravo, and P. R. Schrater. Effect of contrast and period on perceived coherence of moving square-wave plaids. In *Investigative Opthalmology and Visual Science Supplement (ARVO)*, volume 36, pages S–51, May 1995.

[22] R. Dubner and S. M. Zeki. Response properties and receptive fields of cells in an anatomically defined region of the superior temporal sulcus in the monkey. *Brain Res*, 35(2):528–532, Dec 1971.

[23] J. M. Allman and J. H. Kaas. A representation of the visual field in the caudal third of the middle tempral gyrus of the owl monkey (aotus trivirgatus). *Brain Res*, 31(1):85–105, Aug 1971.

[24] T. D. Albright, R. Desimone, and C. G. Gross. Columnar organization of directionally selective cells in visual area mt of the macaque. *J Neurophysiol*, 51(1):16–31, Jan 1984.

[25] S. G. Lisberger, E. J. Morris, and L. Tychsen. Visual motion processing and sensory-motor integration for smooth pursuit eye movements. *Annu Rev Neurosci*, 10:97–129, 1987.

[26] S. M. Zeki. Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *J Physiol*, 236(3):549–573, Feb 1974.

[27] S. Zeki. The response properties of cells in the middle temporal area (area mt) of owl monkey visual cortex. *Proc R Soc Lond B Biol Sci*, 207(1167):239–248, Feb 1980.

[28] J. F. Baker, S. E. Petersen, W. T. Newsome, and J. M. Allman. Visual response properties of neurons in four extrastriate visual areas of the owl monkey (aotus trivirgatus): a quantitative comparison of medial, dorsomedial, dorsolateral, and middle temporal areas. *J Neurophysiol*, 45(3):397–416, Mar 1981.

[29] D. C. Van Essen, J. H. Maunsell, and J. L. Bixby. The middle temporal visual area in the macaque: myeloarchitecture, connections, functional properties and topographic organization. *J Comp Neurol*, 199(3):293–326, Jul 1981.

[30] J. H. Maunsell and D. C. Van Essen. Functional properties of neurons in middle temporal visual area of the macaque monkey. ii. binocular interactions

and sensitivity to binocular disparity. *J Neurophysiol*, 49(5):1148–1167, May 1983.

[31] J. H. Maunsell and D. C. Van Essen. Functional properties of neurons in middle temporal visual area of the macaque monkey. i. selectivity for stimulus direction, speed, and orientation. *J Neurophysiol*, 49(5):1127–1147, May 1983.

[32] D. J. Felleman and J. H. Kaas. Receptive-field properties of neurons in middle temporal visual area (mt) of owl monkeys. *J Neurophysiol*, 52(3):488–513, Sep 1984.

[33] G. A. Orban. Visual processing in macaque area MT/V5 and its satellites (MSTd and MSTv). *Cerebral Cortex*, pages 359–434, 1997.

[34] K. H. Britten. The middle temporal area: Motion processing and the link to perception. *The visual neurosciences*, 2:1203–1216, 2003.

[35] R. T. Born and D. C. Bradley. Structure and function of visual area mt. *Annu Rev Neurosci*, 28:157–189, 2005.

[36] H. R. Rodman and T. D. Albright. Coding of visual stimulus velocity in area mt of the macaque. *Vision Res*, 27(12):2035–2048, 1987.

[37] N. C. Rust, V. Mante, E. P. Simoncelli, and J. A. Movshon. How mt cells analyze the motion of visual patterns. *Nat Neurosci*, 9(11):1421–1431, Nov 2006.

[38] J. A. Movshon and W. T. Newsome. Visual response properties of striate cortical neurons projecting to area mt in macaque monkeys. *J Neurosci*, 16(23):7733–7741, Dec 1996.

[39] G. A. Orban, H. Kennedy, and J. Bullier. Velocity sensitivity and direction selectivity of neurons in areas v1 and v2 of the monkey: influence of eccentricity. *J Neurophysiol*, 56(2):462–480, Aug 1986.

[40] S. J. Prince, A. D. Pointon, B. G. Cumming, and A. J. Parker. The precision of single neuron responses in cortical area v1 during stereoscopic depth judgments. *J Neurosci*, 20(9):3387–3400, May 2000.

[41] R. Gattass and C. G. Gross. Visual topography of striate projection zone (mt) in posterior superior temporal sulcus of the macaque. *J Neurophysiol*, 46(3):621–638, Sep 1981.

[42] M. M. Churchland, N. J. Priebe, and S. G. Lisberger. Comparison of the spatial limits on direction selectivity in visual areas mt and v1. *J Neurophysiol*, 93(3):1235–1245, Mar 2005.

[43] K. Nakayama. Biological image motion processing: a review. *Vision Res*, 25(5):625–660, 1985.

[44] P. Cavanagh and G. Mather. Motion: the long and short of it. *Spat Vis*, 4(2-3):103–129, 1989.

[45] C. Chubb and G. Sperling. Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proc Natl Acad Sci U S A*, 86(8):2985–2989, Apr 1989.

[46] M. N. Shadlen, E. Zohary, K. H. Britten, and W. T. Newsome. Directional properties of MT neurons examined with motion energy filtered apparent motion stimuli. *Investigative Ophthalmology and Visual Science (Suppl.)*, 34:1027, 1993.

[47] W. T. Newsome and E. B. Pare. A selective impairment of motion perception following lesions of the middle temporal visual area (mt). *J Neurosci*, 8(6):2201–2211, Jun 1988.

[48] W. T. Newsome, K. H. Britten, and J. A. Movshon. Neuronal correlates of a perceptual decision. *Nature*, 341(6237):52–54, Sep 1989.

[49] K. H. Britten, M. N. Shadlen, W. T. Newsome, and J. A. Movshon. The analysis of visual motion: a comparison of neuronal and psychophysical performance. *J Neurosci*, 12(12):4745–4765, Dec 1992.

[50] K. H. Britten, W. T. Newsome, M. N. Shadlen, S. Celebrini, and J. A. Movshon. A relationship between behavioral choice and the visual responses of neurons in macaque mt. *Vis Neurosci*, 13(1):87–100, 1996.

[51] M. N. Shadlen, K. H. Britten, W. T. Newsome, and J. A. Movshon. A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *J Neurosci*, 16(4):1486–1510, Feb 1996.

[52] C. D. Salzman, K. H. Britten, and W. T. Newsome. Cortical microstimulation influences perceptual judgements of motion direction. *Nature*, 346(6280):174–177, Jul 1990.

[53] C. D. Salzman, C. M. Murasugi, K. H. Britten, and W. T. Newsome. Microstimulation in visual area mt: effects on direction discrimination performance. *J Neurosci*, 12(6):2331–2355, Jun 1992.

[54] C. D. Salzman and W. T. Newsome. Neural mechanisms for forming a perceptual decision. *Science*, 264(5156):231–237, Apr 1994.

[55] A. B. Watson and A. J. Ahumada. A look at motion in the frequency domain. *Motion: Perception and representation*, pages 1–10, 1983.

[56] R. C. Reid, R. E. Soodak, and R. M. Shapley. Linear mechanisms of directional selectivity in simple cells of cat striate cortex. *Proc Natl Acad Sci U S A*, 84(23):8740–8744, Dec 1987.

[57] J. McLean and L. A. Palmer. Contribution of linear spatiotemporal receptive field structure to velocity selectivity of simple cells in area 17 of cat. *Vision Res*, 29(6):675–679, 1989.

[58] J. A. Movshon, I. D. Thompson, and D. J. Tolhurst. Receptive field organization of complex cells in the cat's striate cortex. *J Physiol*, 283:79–99, Oct 1978.

[59] E. H. Adelson and J. R. Bergen. Spatiotemporal energy models for the perception of motion. *J Opt Soc Am A*, 2(2):284–299, Feb 1985.

[60] D. J. Heeger. Model for the extraction of image flow. *J Opt Soc Am A*, 4(8):1455–1471, Aug 1987.

[61] N. M. Grzywacz and A. L. Yuille. A model for the estimate of local image velocity by cells in the visual cortex. *Proc R Soc Lond B Biol Sci*, 239(1295):129–161, Mar 1990.

[62] E. P. Simoncelli and D. J. Heeger. A model of neuronal responses in visual area mt. *Vision Res*, 38(5):743–761, Mar 1998.

[63] E. P. Simoncelli, W. D. Bair, J. R. Cavanaugh, and J. A. Movshon. Testing and refining a computational model of neural responses in area MT. In *Investigative Opthalmology and Visual Science Supplement (ARVO)*, volume 37, pages S–916, May 1996.

[64] P. R. Schrater, D. C. Knill, and E. P. Simoncelli. Mechanisms of visual motion detection. *Nat Neurosci*, 3(1):64–68, Jan 2000.

[65] C. Blakemore, J. Nachmias, and P. Sutton. The perceived spatial frequency shift: evidence for frequency-selective neurones in the human brain. *J Physiol*, 210(3):727–750, Oct 1970.

[66] C. Blakemore and P. Sutton. Size adaptation: a new aftereffect. *Science*, 166(902):245–247, Oct 1969.

[67] A. Pantle and R. Sekuler. Size-detecting mechanisms in human vision. *Science*, 162(858):1146–1148, Dec 1968.

[68] E. Levinson and R. Sekuler. Adaptation alters perceived direction of motion. *Vision Res*, 16(7):779–781, 1976.

[69] E. Levinson and R. Sekuler. A two-dimensional analysis of direction-specific adaptation. *Vision Res*, 20(2):103–107, 1980.

[70] F. W. Campbell and L. Maffei. The tilt after-effect- A fresh look(Neurophysiological investigation of visual tilt aftereffect, comparing judgment precision at vertical and horizontal to oblique orientation with/without gravity cue). *Vision Research*, 11:833–840, 1971.

[71] R. H. Carpenter and C. Blakemore. Interactions between orientations in human vision. *Exp Brain Res*, 18(3):287–303, Oct 1973.

[72] J. E. Calvert and J. P. Harris. Spatial frequency and duration effects on the tilt illusion and orientation acuity. *Vision Res*, 28(9):1051–1059, 1988.

[73] C. W. Clifford, P. Wenderoth, and B. Spehar. A functional angle on some after-effects in cortical vision. *Proc Biol Sci*, 267(1454):1705–1710, Sep 2000.

[74] O. Schwartz, A. Hsu, and P. Dayan. Space and time in visual context. *Nat Rev Neurosci*, 8(7):522–535, Jul 2007.

[75] L. Maffei, A. Fiorentini, and S. Bisti. Neural correlate of perceptual adaptation to gratings. *Science*, 182(116):1036–1038, Dec 1973.

[76] M. J. Wright and A. Johnston. Invariant tuning of motion aftereffect. *Vision Res*, 25(12):1947–1955, 1985.

[77] S. Nishida and A. Johnston. Influence of motion signals on the perceived position of spatial pattern. *Nature*, 397(6720):610–612, Feb 1999.

[78] E. Hiris and R. Blake. Direction repulsion in motion transparency. *Vis Neurosci*, 13(1):187–197, 1996.

[79] P. Thompson. Velocity after-effects: the effects of adaptation to moving stimuli on the perception of subsequently seen moving stimuli. *Vision Res*, 21(3):337–345, 1981.

[80] A. T. Smith. Velocity coding: evidence from perceived velocity shifts. *Vision Res*, 25(12):1969–1976, 1985.

[81] A. T. Smith and G. K. Edgar. Antagonistic comparison of temporal frequency filter outputs as a basis for speed perception. *Vision Research*, 34(2):253, 1994.

[82] T. Ledgeway and A. T. Smith. Changes in perceived speed following adaptation to first-order and second-order motion. *Vision Res*, 37(2):215–224, Jan 1997.

[83] P. R. Schrater and E. P. Simoncelli. Local velocity representation: evidence from motion adaptation. *Vision Res*, 38(24):3899–3912, Dec 1998.

[84] G. Mather, F. Verstraten, and S. M. Anstis. *The motion aftereffect: A modern perspective*. Mit Press, 1998.

[85] C. W. Clifford and P. Wenderoth. Adaptation to temporal modulation can enhance differential speed sensitivity. *Vision Res*, 39(26):4324–4332, Oct 1999.

[86] A. A. Stocker and E. P. Simoncelli. Visual motion aftereffects arise from a cascade of two isomorphic adaptation mechanisms. *Journal of Vision*, 9(9):1–14, Aug 2009.

[87] F. A. Verstraten, R. Verlinde, R. E. Fredericksen, and W. A. van de Grind. A transparent motion aftereffect contingent on binocular disparity. *Perception*, 23(10):1181–1188, 1994.

[88] F. A. Verstraten, M. J. van der Smagt, R. E. Fredericksen, and W. A. van de Grind. Integration after adaptation to transparent motion: static and dynamic test patterns result in different aftereffect directions. *Vision Res*, 39(4):803–810, Feb 1999.

[89] C. P. Benton and W. Curran. Direction repulsion goes global. *Curr Biol*, 13(9):767–771, Apr 2003.

[90] M. von Grunau. Bivectorial transparent stimuli simultaneously adapt mechanisms at different levels of the motion pathway. *Vision Res*, 42(5):577–587, Mar 2002.

[91] R. J. Snowden. Adaptability of the visual system is inversely related to its sensitivity. *J Opt Soc Am A Opt Image Sci Vis*, 11(1):25–32, Jan 1994.

[92] A. Clymer. *The effects of seen motion on the apparent speed of subsequent test velocities: speed tuning of movement aftereffects*. PhD thesis, Ph. D. Thesis, Columbia University, New York, 1973.

[93] P. Thompson. Perceived rate of movement depends on contrast. *Vision Res*, 22(3):377–380, 1982.

[94] E. P. Simoncelli, E. H. Adelson, and D. J. Heeger. Probability distributions of optical flow. In *Proc Conf on Computer Vision and Pattern Recognition*, pages 310–315, Mauii, Hawaii, Jun 3-6 1991. IEEE Computer Society. Original version: MIT Media Lab TR.

[95] E. P. Simoncelli. *Distributed Analysis and Representation of Visual Motion.* PhD thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, January 1993.

[96] Y. Weiss. *Bayesian motion estimation and segmentation.* PhD thesis, Massachusetts Institute of Technology, May 1998.

[97] E. Koechlin, J. L. Anton, and Y. Burnod. Bayesian inference in populations of cortical neurons: a model of motion integration and segmentation in area mt. *Biol Cybern*, 80(1):25–44, Jan 1999.

[98] D. Ascher and N. M. Grzywacz. A bayesian model for the measurement of visual velocity. *Vision Res*, 40(24):3427–3434, 2000.

[99] Y. Weiss and D. J. Fleet. *Velocity likelihoods in biological and machine vision. Probabilistic models of the brain: Perception and neural function.* Cambridge, MA: MIT Press, 2001.

[100] Y. Weiss, E. P. Simoncelli, and E. H. Adelson. Motion illusions as optimal percepts. *Nat Neurosci*, 5(6):598–604, Jun 2002.

[101] P. Thompson, L. S. Stone, and S. Swash. Speed estimates from grating patches are not contrast-normalized. *Vision Res*, 36(5):667–674, Mar 1996.

[102] L. S. Stone, A. B. Watson, and J. B. Mulligan. Effect of contrast on the perceived direction of a moving plaid. *Vision Res*, 30(7):1049–1067, 1990.

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

[104] A. A. Stocker and E. P. Simoncelli. Noise characteristics and prior expectations in human visual speed perception. *Nat Neurosci*, 9(4):578–585, Apr 2006.

[105] D. H. Hubel and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *J Physiol*, 160:106–154, Jan 1962.

[106] D. H. Hubel and T. N. Wiesel. Receptive fields and functional architecture of monkey striate cortex. *J Physiol*, 195(1):215–243, Mar 1968.

[107] I. Fujita, K. Tanaka, M. Ito, and K. Cheng. Columns for visual features of objects in monkey inferotemporal cortex. *Nature*, 360(6402):343–346, Nov 1992.

[108] J. L. Gallant, J. Braun, and D. C. Van Essen. Selectivity for polar, hyperbolic, and cartesian gratings in macaque visual cortex. *Science*, 259(5091):100–103, Jan 1993.

[109] E. Kobatake and K. Tanaka. Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *J Neurophysiol*, 71(3):856–867, Mar 1994.

[110] F. Wilkinson, T. W. James, H. R. Wilson, J. S. Gati, R. S. Menon, and M. A. Goodale. An fmri study of the selective activation of human extrastriate form vision areas by radial and concentric gratings. *Curr Biol*, 10(22):1455–1458, Nov 2000.

[111] D. J. Heeger, E. P. Simoncelli, and J. A. Movshon. Computational models of cortical visual processing. *Proc Natl Acad Sci U S A*, 93(2):623–627, Jan 1996.

[112] A. Thiele, K. R. Dobkins, and T. D. Albright. Neural correlates of contrast detection at threshold. *Neuron*, 26(3):715–724, Jun 2000.

[113] M. J. Hawken, A. J. Parker, and J. S. Lund. Laminar organization and contrast sensitivity of direction-selective cells in the striate cortex of the old world monkey. *J Neurosci*, 8(10):3541–3548, Oct 1988.

[114] T. D. Albright. Direction and orientation selectivity of neurons in visual area mt of the macaque. *J Neurophysiol*, 52(6):1106–1130, Dec 1984.

[115] H. R. Rodman and T. D. Albright. Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Exp Brain Res*, 75(1):53–64, 1989.

[116] T. D. Albright and R. Desimone. Local precision of visuotopic organization in the middle temporal area (mt) of the macaque. *Exp Brain Res*, 65(3):582–592, 1987.

[117] J. R. Cavanaugh, W. Bair, and J. A. Movshon. Selectivity and spatial distribution of signals from the receptive field surround in macaque v1 neurons. *J Neurophysiol*, 88(5):2547–2556, Nov 2002.

[118] K. Tanaka, K. Hikosaka, H. Saito, M. Yukie, Y. Fukada, and E. Iwai. Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *J Neurosci*, 6(1):134–144, Jan 1986.

[119] K. H. Britten and H. W. Heuer. Spatial summation in the receptive fields of mt neurons. *J Neurosci*, 19(12):5074–5084, Jun 1999.

[120] A. Mikami, W. T. Newsome, and R. H. Wurtz. Motion selectivity in macaque visual cortex. ii. spatiotemporal range of directional interactions in mt and v1. *J Neurophysiol*, 55(6):1328–1339, Jun 1986.

[121] W. T. Newsome, A. Mikami, and R. H. Wurtz. Motion selectivity in macaque visual cortex. iii. psychophysics and physiology of apparent motion. *J Neurophysiol*, 55(6):1340–1351, Jun 1986.

[122] A. Mikami, W. T. Newsome, and R. H. Wurtz. Motion selectivity in macaque visual cortex. i. mechanisms of direction and speed selectivity in extrastriate area mt. *J Neurophysiol*, 55(6):1308–1327, Jun 1986.

[123] J. A. Perrone and A. Thiele. A model of speed tuning in mt neurons. *Vision Res*, 42(8):1035–1051, Apr 2002.

[124] O. Braddick. A short-range process in apparent motion. *Vision Res*, 14(7):519–527, Jul 1974.

[125] C. L. Baker and O. J. Braddick. Eccentricity-dependent scaling of the limits for short-range apparent motion perception. *Vision Res*, 25(6):803–812, 1985.

[126] U. J. Ilg and J. Churan. Motion perception without explicit activity in areas mt and mst. *J Neurophysiol*, 92(3):1512–1523, Sep 2004.

[127] O. J. Braddick. Low-level and high-level processes in apparent motion. *Philos Trans R Soc Lond B Biol Sci*, 290(1038):137–151, Jul 1980.

[128] N. J. Majaj, M. Carandini, and J. A. Movshon. Motion integration by neurons in macaque mt is local, not global. *J Neurosci*, 27(2):366–370, Jan 2007.

[129] S. Wuerger, R. Shapley, and N. Rubin. "On the visually perceived direction of motion"by Hans Wallach: 60 years later. *Perception*, 25:1317–1367, 1996.

[130] P. Burt and G. Sperling. Time, distance, and feature trade-offs in visual apparent motion. *Psychol Rev*, 88(2):171–195, Mar 1981.

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

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

[133] L. S. Stone and P. Thompson. Human speed perception is contrast dependent. *Vision Res*, 32(8):1535–1549, Aug 1992.

[134] O. Braddick. Segmentation versus integration in visual motion processing. *Trends Neurosci*, 16(7):263–268, Jul 1993.

[135] S. J. Cropper, K. T. Mullen, and D. R. Badcock. Motion coherence across different chromatic axes. *Vision Res*, 36(16):2475–2488, Aug 1996.

[136] F. L. Kooi, K. K. De Valois, E. Switkes, and D. H. Grosof. Higher-order factors influencing the perception of sliding and coherence of a plaid. *Perception*, 21(5):583–598, 1992.

[137] J. Krauskopf and B. Farell. Influence of colour on the perception of coherent motion. *Nature*, 348(6299):328–331, Nov 1990.

[138] J. M. Hupe and N. Rubin. The dynamics of bi-stable alternation in ambiguous motion displays: a fresh look at plaids. *Vision Res*, 43(5):531–548, Mar 2003.

[139] T. J. Andrews and D. Schluppeck. Ambiguity in the perception of moving stimuli is resolved in favour of the cardinal axes. *Vision Res*, 40(25):3485–3493, 2000.

[140] F. A. Wichmann and N. J. Hill. The psychometric function: I. Fitting, sampling, and goodness of fit. *Perception and Psychophysics*, 63(8):1293–1313, 2001.

[141] F. Hurlimann, D. C. Kiper, and M. Carandini. Testing the bayesian model of perceived speed. *Vision Res*, 42(19):2253–2257, Sep 2002.

[142] P. Thompson, K. Brooks, and S. T. Hammett. Speed can go up as well as down at low contrast: implications for models of motion perception. *Vision Res*, 46(6-7):782–786, Mar 2006.

[143] W. Marshak and R. Sekuler. Mutual repulsion between moving visual targets. *Science*, 205(4413):1399–1401, Sep 1979.

[144] O. J. Braddick, K. A. Wishart, and W. Curran. Directional performance in motion transparency. *Vision Res*, 42(10):1237–1248, May 2002.

[145] D. Burke and P. Wenderoth. The effect of interactions between one-dimensional component gratings on two-dimensional motion perception. *Vision Res*, 33(3):343–350, Feb 1993.

[146] J. Kim and H. R. Wilson. Direction repulsion between components in motion transparency. *Vision Res*, 36(8):1177–1187, Apr 1996.

[147] H. J. Rauber and S. Treue. Revisiting motion repulsion: evidence for a general phenomenon? *Vision Res*, 39(19):3187–3196, Sep 1999.

[148] E. C. Hildreth. *Measurement of Visual Motion*. MIT Press, Cambridge, MA, USA, 1984.

[149] E. P. Simoncelli. Local analysis of visual motion. In L. M. Chalupa and J. S. Werner, editors, *The Visual Neurosciences*, chapter 109, pages 1616–1623. MIT Press, January 2003.

[150] M. von Grunau and S. Dube. Comparing local and remote motion aftereffects. *Spat Vis*, 6(4):303–314, 1992.

[151] F. A. Verstraten, R. E. Fredericksen, and W. A. van de Grind. Movement aftereffect of bi-vectorial transparent motion. *Vision Res*, 34(3):349–358, Feb 1994.

[152] F. A. Verstraten, R. E. Fredericksen, R. J. van Wezel, J. C. Boulton, and W. A. van de Grind. Directional motion sensitivity under transparent motion conditions. *Vision Res*, 36(15):2333–2336, Aug 1996.

[153] F. A. Verstraten, R. E. Fredericksen, R. J. Van Wezel, M. J. Lankheet, and W. A. Van de Grind. Recovery from adaptation for dynamic and static motion aftereffects: evidence for two mechanisms. *Vision Res*, 36(3):421–424, Feb 1996.

[154] A. Kohn and J. A. Movshon. Adaptation changes the direction tuning of macaque mt neurons. *Nat Neurosci*, 7(7):764–772, Jul 2004.