Title
A Computational Theory for Sensory Adaptation

Author
Lee, Alan Lap Fai

Publication Date
2013

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Los Angeles

A Computational Theory for Sensory Adaptation

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Psychology

by

Alan Lap Fai Lee

2013

ABSTRACT OF THE DISSERTATION

A Computational Theory

for Sensory Adaptation

by

Alan Lap Fai Lee

Doctor of Philosophy in Psychology

University of California, Los Angeles, 2013

Professor Hongjing Lu, Chair

Our sensory system consists of multiple processing stages, and its response characteristics change based on recent stimulus history. Although much is known about both the hierarchical nature and the adaptability of the sensory system, it remains unclear how this multilevel neural system adapts to changes in the environment and leads to various perceptual consequences. In my dissertation, I focus on adaptation of the visual motion system. I aim to address the following questions: (I) Can the adaptation of the hierarchical motion system be studied at individual levels of processing? (II) How does the hierarchical motion system adapt and produce perceptual aftereffects? (III) What is the computational principle that underlies multilevel adaptation in the

motion system? I designed a novel psychophysical paradigm to study how adaptation of individual levels of motion processing may contribute in producing various forms of perceptual aftereffects. My psychophysical findings suggest that perceptual aftereffects depend on the adaptation of multiple stages of processing, each may adapt based on its own computational principle. Finally, I used a multilayer network model to simultaneously explain different adaptation-induced neural changes that have been observed at different processing levels. Differences in neural adaptation at the local and the global levels of motion processing can be explained from a natural-statistics perceptive: neurons at different levels are holding different assumptions about natural sensory statistics, causing them to react differently in response to the same adapting environment. Taken together, findings from this dissertation illustrate the importance of understanding sensory adaptation from a hierarchical-processing perspective, which will shed more lights on the big picture of sensory processing.

The dissertation of Alan Lap Fai Lee is approved.

James W. Bisley

Dario L. Ringach

Yingnian Wu

Alan L. Yuille

Hongjing Lu, Committee Chair

University of California, Los Angeles

2013

Table of Contents

Lists of Figures

List of Table

Acknowledgements

This research was supported by the following:

- UCLA Distinguished University Fellowship
- UCLA Department of Psychology Excellence in Research Award
- UCLA Faculty Research Grant (Principal Investigator: Hongjing Lu)
  Title: Phantom motion aftereffects in global motion perception
- National Science Foundation Grant (BCS-0843880; Principal Investigator: Hongjing Lu)
  Title: CAREER: A Computational Investigation into Biological Motion Perception
- Elsevier/Vision Research Travel Award
  for the 2012 Annual Vision Sciences Society Meeting
- UCLA Dissertation Year Fellowship

Two chapters in this dissertation were published on *Journal of Vision*, an open-access journal:

Chapter 2: Lee, A. L. F., & Lu, H. (2010). A comparison of global motion perception using a multiple-aperture stimulus. Journal of Vision, 10(4).
Chapter 3: Lee, A. L. F., & Lu, H. (2012). Two forms of aftereffects induced by transparent motion reveal multilevel adaptation. Journal of Vision, 12(4).

I am deeply indebted to my advisor, Hongjing Lu, for her guidance, advice and support throughout my five years at UCLA. What's influenced me is not only her interest and enthusiasm about science, but also her attitude towards many different things in life. I would also like to thank other members of my committee (James Bisley, Dario Ringach, Yingnian Wu and Alan Yuille), who have given me invaluable advice at various points in my graduate years.

Other faculty members that I'd like to thank include Ladan Shams, Zili Liu, and Phil Kellman, who have given me helpful feedback and advice on my work. Many thanks to Hakwan Lau, for inspiring me to look at science from a totally different perspective, which has made it more challenging, exciting and meaningful than how I had looked at it before.

There are also many friends I would like to thank. My labmates: Jeroen, Junzhu, and Steve, who have helped me a lot in both research and many other things; my fellow Psych Grads: especially Brian, Chris, Dawn, Everett, Genna, KP, Megan and Mike Friedman, for the fun times of TAing/meetings/conferences we spent together; and most importantly, Lindsay, for her assistance and all the important reminders.

Lastly, I would like to thank my family and friends in Hong Kong, for giving me tremendous support throughout the years. Dad, Albert, Bonnie, and Alfred: Thanks for taking care of me. I am really grateful to be part of this wonderful family. I am sure mom would be proud if she was here. Thanks to all my uncles and aunts, who have looked after me for so many years. Many thanks to my friends with whom I've discussed my research and career, especially my fellow CogSci classmates, including Anton, Elaine, Joey, Kin-Chi, Louis & Ophe. Thanks to Jenny, for the support during many important events in my life. And, finally, to Jodie: This final year has been quite hectic, but thanks for making it a much more colorful and enjoyable one. I'm grateful for the changes you've brought about in my life.

Vita

2001 – 2005	B.S. in Cognitive Science (First Class Honours) University of Hong Kong (HKU)
2005 – 2007	Management Trainee HSBC (Hong Kong)
2007 – 2008	M.Phil. in Psychology (incomplete) University of Hong Kong
2008 – 2010	M.A. in Psychology University of California, Los Angeles (UCLA)
2008 – 2012	Teaching Assistant Department of Psychology University of California, Los Angeles
2012	*Elsevier/Vision Research* Travel Award for the 2012 Annual Vision Sciences Society (VSS) Meeting
2012 – 2013	UCLA Dissertation Year Fellowship

PUBLICATIONS

- **Lee, A. L. F.**, & Lu, H. (2012). Two forms of aftereffects induced by transparent motion reveal multilevel adaptation. *Journal of Vision*, *12*(4):3, 1–13.
- **Lee, A. L. F.**, & Lu, H. (2010). A comparison of global motion perception using a multiple-aperture stimulus. *Journal of Vision*, *10*(4): 9, 1-16.
- Lu, H., Lin, T., **Lee, A. L. F.,** Vese, L., & Yuille, A. (2010). Functional form of motion priors in human motion perception. *Advances in Neural Information Processing Systems*, *23*, 1495-1503. Cambridge, MA: MIT Press.
- Wu, S., Lu, H., **Lee, A.**, & Yuille A. "Motion integration using competitive priors", In: *Statistical and Geometrical Approaches to Visual Motion Analysis*, Cremers, D., Rosenhahn, B., Yuille, A.L., & Schmidt, F.R. (Eds.), Springer, 5604:1 - 26 (2009)

CONFERENCE PRESENTATIONS (*Oral presentations)

1. ***Lee, A. L. F.**, Lu, H. (2013). Propagation of local adaptation is insufficient to generate repulsive motion aftereffects. *VSS*, Naples, Florida.
2. *Lu, H., **Lee, A. L. F.**, (2013). Inferring "hidden" parts by learning hierarchical representations of objects. *VSS*, Naples, Florida.
3. ***Lee, A. L. F.**, Lu, H. (2012). Adapting to imperceptible multidirectional motion yields perceptible aftereffects: A psychophysical and computational investigation. *VSS*, Naples, Florida.
4. ***Lee, A. L. F.**, Lu, H. (2011). A new form of motion aftereffect in transparent motion adaptation. *VSS*, Naples, Florida.
5. *Lu, H., Lin, T., **Lee, A. L. F.**, Vese, L., & Yuille, A. (2010). Recovering the functional form of the slowand-smooth prior in global motion perception. *VSS*, Naples, Florida.
6. **Lee, A. L. F.** & Lu, H. (2010). Phantom motion aftereffect using multiple aperture stimuli: A dynamic Bayesian model. *VSS*, Naples, Florida.
7. Lu, H., **Lee, A.**, & Keane, B. P. (2009). Spatial pattern analysis in biological motion. *VSS*, Naples, Florida.
8. **Lee, A.** & Lu, H. (2009). A comparison of motion integration for optic flow components. *VSS*, Naples, Florida.
9. **Lee, A. L.**, Yuille, A., & Lu, H. (2008). Superior perception of circular/radial than translational motion cannot be explained by generic priors. *VSS*, Naples, Florida.
10. ***Lee, A. L.**, Yuille, A., & Lu, H. (2008). Superior perception of circular/radial over translational motion cannot be explained by generic priors. *Hong Kong Vision Meeting*.

Chapter 1

Introduction

Visual adaptation produces remarkable perceptual aftereffects. However, it remains unclear what basic neural mechanisms underlie visual adaptation, and how these adaptation-induced neural changes are related to perceptual aftereffects. Previous neurophysiological studies have found that, after prolonged stimulation, neurons in the visual cortex change their response characteristics (Clark & Bradley, 2008; Dragoi, Sharma, & Sur, 2000; Kohn & Movshon, 2003, 2004; Krekelberg, van Wezel, & Albright, 2006), and that such adaptation-induced neural changes are related to remarkable perceptual aftereffects (C. W. Clifford, et al., 2007; Schrater & Simoncelli, 1998; F. A. Verstraten, Fredericksen, & van de Grind, 1994). The special case of contrast adaptation has been extensively studied over the past 20 years. Converging physiological and psychophysical evidence has shown that contrast adaptation induces neural changes through two mechanisms. First, adaptation independently modulates neural activity at early processing stages, as early as in retina and LGN (Baccus & Meister, 2004; Solomon, Peirce, Dhruv, & Lennie, 2004); second, these low-level changes are propagated up the visual hierarchy to affect neural responses in higher-level areas such as MT (Kohn & Movshon, 2003). It remains unclear whether these two basic mechanisms are also recruited for other types of sensory adaptation, and whether adaptation-induced neural changes at different processing levels can lead to distinct perceptual aftereffects. My dissertation will study motion adaptation to provide behavioral examination and computational accounts for the basic question: How does the hierarchical sensory system adapt? I will first review two research topics that have been

extensively studied in isolation over decades, but have not been well understood in relation to one another. I will then show how my thesis will connect the two topics and provide a general framework for adaptation in motion hierarchy.

The sensory system is hierarchical.

Over the past several decades, research from various disciplines has provided converging evidence to support the hypothesis that sensory processing involves multiple stages. Generally speaking, at the early stages, location-specific sensory signals are extracted. These local signals are then integrated in later stages of processing to produce a global interpretation or percept of the world.

For example, in the case of visual motion processing, evidence from multiple perspectives (Britten, Shadlen, Newsome, & Movshon, 1993; Duffy & Wurtz, 1991a, 1991b; Emerson, Bergen, & Adelson, 1992; Morrone, Burr, & Vaina, 1995; Qian & Andersen, 1994; Qian, Andersen, & Adelson, 1994a, 1994b; Rust, Mante, Simoncelli, & Movshon, 2006; Snowden & Milne, 1997) supports the conclusion that the motion system involves multiple stages of processing. Local motion signals are first extracted by neurons in the primary visual cortex (V1). These local signals are ambiguous in representing the underlying global motion direction (the "aperture problem", Marr & Ullman, 1981), and need to be integrated to produce a percept or “solution” of the global motion. It has been shown that these global motion signals specifically activate neurons in the middle temporal (MT) area in the primate brain, and are correlated with activity in the human MT complex (hMT+; Amano, et al., 2012; Huk & Heeger, 2002; Tootell, et al., 1995). Furthermore, motion patterns that are more complex, such as circular or radial

motion, are analyzed by higher-level mechanisms and cortical areas (Duffy & Wurtz, 1991a, 1991b; Tanaka, Fukada, & Saito, 1989; Tanaka & Saito, 1989; Wall & Smith, 2008).

The sensory system is adaptive

The sensory system is not a rigid machine that remains unchanged across different situations. Rather, it changes its response characteristics in response to recent stimulus history. Prolonged exposure to a certain stimulus, or sensory adaptation, causes robust and systematic changes in the system. These adaptation-induced effects can be observed and analyzed using various methods, including single-cell recording, neuroimaging, psychophysics and computational models
For example, prolonged exposure to a motion stimulus leads to strong, illusory perceptual biases, collectively known as motion aftereffects (MAE; Anstis, Verstraten, & Mather, 1998; Mather, Pavan, Campana, & Casco, 2008). Perceived motion direction can also be systematically biased after adapting to a certain motion direction (Levinson & Sekuler, 1976; Schrater & Simoncelli, 1998). In neurophysiology, it has been shown that prolonged stimulation to sensory neurons also leads to changes in neural tuning (Dragoi, et al., 2000; Kohn & Movshon, 2004; Krekelberg, et al., 2006), which is believed to be the cause of perceptual aftereffects (Jin, Dragoi, Sur, & Seung, 2005; Schwartz, Hsu, & Dayan, 2007).

Hierarchical adaptation in the motion system

Even though both the hierarchical nature and the adaptability of the sensory system have been extensively studied, little is known about how the system adapts as a hierarchy. Indeed, a few recent reviews have pointed out the importance of understanding sensory adaptation in the context of hierarchical processing (C. W. Clifford, et al., 2007; Kohn, 2007). In the past decade,

this direction has drawn the attention of researchers from many fields, particular in vision science, in which effects of adaptation can be traced from as early as the retina to the lateral geniculate nuclei (LGN) (Solomon, et al., 2004), from LGN to V1 (A. A. Stocker & Simoncelli, 2009), from V1 to MT (Curran, Clifford, & Benton, 2006; Kohn & Movshon, 2003, 2004), and from V1 to face-sensitive areas such as IT (Xu, Dayan, Lipkin, & Qian, 2008). By viewing adaptation as a process of stimulus-induced changes that occur at distinct processing levels, one can ask many interesting and challenging questions regarding the nature, the process and the computation. My dissertation aims to address the following questions: (1) Is it possible to delineate the effects of adaptation at individual processing levels in a hierarchical system? (2) How does adaptation in a multilevel neural system produce perceptual aftereffects? (3) What computational principles underlie an adaptive hierarchy of sensory processing?

Dissertation overview

My dissertation aims to address the above questions related to hierarchical adaptation in the motion processing system. This dissertation is divided into seven chapters. After Chapter 1 (the present Introduction). Chapters 2 and 3 address question (1), aiming to determine whether the adaptation of a hierarchical system can be studied at individual levels of processing. To seek evidence for the existence of higher-level motion processing, psychophysical experiments in Chapter 2 were designed to study high-level integration mechanisms for complex global motion patterns. Chapter 3 reports a few psychophysical experiments probing the local and global levels in the motion processing hierarchy for effects of adaptation.

Chapters 4 and 5 address question (2), concerning how a multilevel neural system adapts and produces perceptual aftereffects. Chapter 4 combines psychophysical evidence and a neural

model to illustrate how motion aftereffects can be generated purely based on propagation of effects resulting from local adaptation. Chapter 5 compares the strength of repulsive direction aftereffect resulting from local and global levels of motion adaptation.

Chapter 6 addresses question (3), which concerns the computational principles that underlie hierarchical adaptation. Chapter 6 presents a multilayer network model that simultaneously captures neural adaptation effects based on low and high levels of motion processing. Chapter 7 summarizes the findings of the dissertation, and their interpretations, and raises some possible directions for future research.

Chapter 2

Integration Mechanisms for Complex Global Motion

Introduction

Imagine that you are traveling on a boat and viewing the moving water through a mesh railing with many holes. Through each hole, you can see the water drifting in different directions. Given these sampled motion signals projected onto the retina, your visual system can readily infer the global motion of the scene, determining, for example, whether the boat is moving forward or backward as well as the speed of the boat's movement. However, this task is far from trivial due to the inherent ambiguity of motion stimuli, usually termed the "aperture problem". As illustrated in Figure 2.1, when only a line segment of a moving object is viewed through a small aperture with the segment's end-points or terminators occluded, there are an infinite number of possible interpretations of the true motion of the object behind this aperture. One way to overcome the local ambiguity inherent in the visual dynamic stimulus is to integrate motion signals viewed from multiple apertures to infer the "true" velocity field for the occluded object.

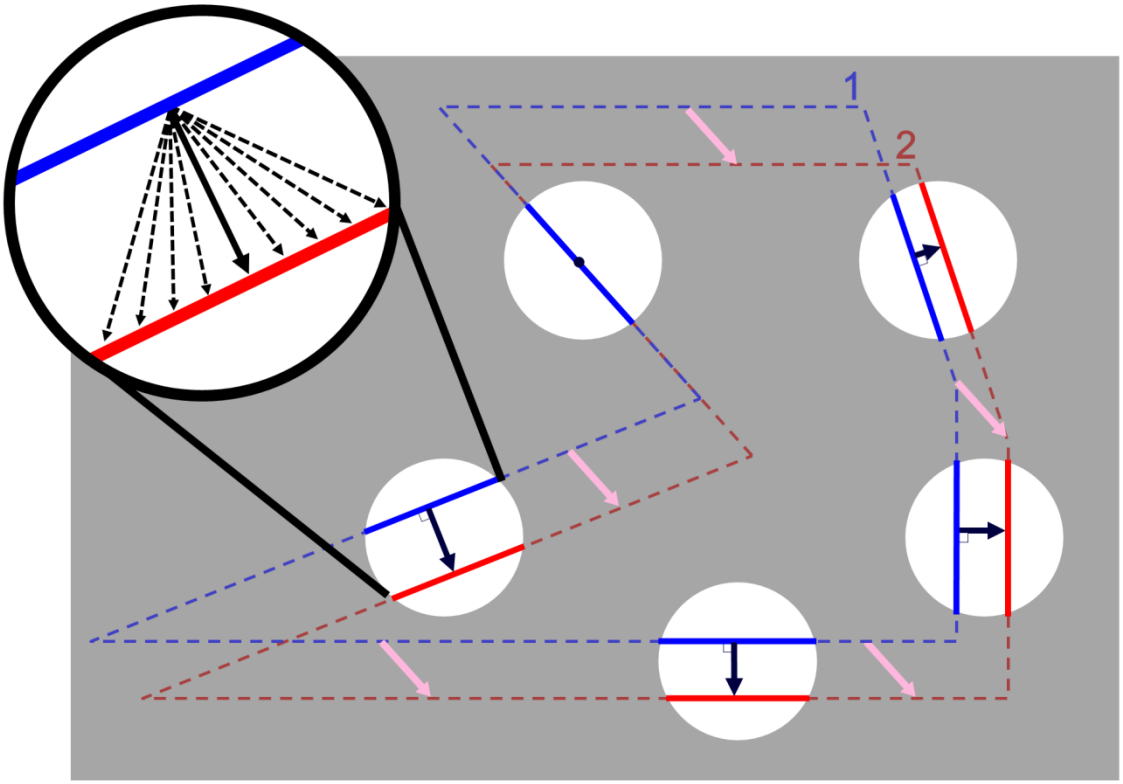

Figure 2.1 Illustration of the aperture problem. As a rigid object moves from position 1 (blue outline) to position 2 (red outline) with a down-rightward motion (pink arrows), each edge, when viewed through an aperture (white circle), is observed to move in the direction (dark blue arrows) orthogonal to its orientation. As shown in the magnified aperture, with only one local motion observed (thick solid arrow), the number of possible interpretations of the true motion of the whole object (thin dashed arrows) is infinite.

Physiological evidence suggests that motion integration depends on a hierarchical motion processing network in the visual cortex. Directional-selective neurons in the primary visual cortex (V1) are sensitive to the motion energy within small receptive fields (Emerson, et al., 1992), which are used to estimate local motion. These local estimates are pooled in higher-level visual areas where neurons with large receptive fields are sensitive to global motion patterns.

These global motion detectors include neurons in the middle temporal (MT) area (Britten, et al., 1993; Rust, et al., 2006) and in the medial superior temporal (MST) area (Duffy & Wurtz, 1991a; Tanaka & Saito, 1989). This hierarchical system for processing motion has been supported by psychophysical studies (Morrone, et al., 1995) and simulated by computational models (Heeger, Simoncelli, & Movshon, 1996).

However, although researchers generally agree that an integration stage is involved in motion processing, it remains unclear whether a common integration mechanism could underlie perception of optic flows viewed in everyday life. Researchers have proposed that the visual system could decompose optic flows into three cardinal components: translational, circular and radial motion (e.g., Burr, Badcock, & Ross, 2001). These three basic motion patterns can jointly represent many kinds of real-world optic flows, generated from moving objects, from saccadic eye movements, and from ego-motion. Accordingly, detectors that are sensitive to these cardinal patterns can work cooperatively to represent a large range of optic flows. Circular and radial motions are often referred to as "complex" motions (Beardsley & Vaina, 2005; Bex, Metha, & Makous, 1999; Burr & Santoro, 2001; C. W. G. Clifford, Beardsley, & Vaina, 1999), because local motion vectors in rotation and expansion change with their locations, whereas local motion vectors in translation are constant over different spatial locations.

A few studies have aimed to compare human sensitivity in perceiving translational, circular and radial motion in order to understand how the visual system conducts a global motion analysis on these motion patterns. However, results for comparisons over the three motion patterns have been inconsistent. The fundamental disagreement focuses on whether the visual system shows similar sensitivity to all three motion patterns or not. Some findings suggest that motion sensitivity for all three patterns is comparable. Blake and Aiba (1998) used high-contrast and

low-density random-dot kinematogram (RDK) to measure motion sensitivity. For both motion detection and direction discrimination tasks, no significant difference was found in motion sensitivity measured with coherence thresholds across the three motion patterns. Other researchers used first-order (luminance-defined) motion stimuli of the RDK type, and found that both contrast and coherence sensitivity for the three motion types were comparable (Aaen-Stockdale, Ledgeway, & Hess, 2007; Bertone & Faubert, 2003). On the other hand, various studies have reported the opposite results, finding that sensitivity differs across translational, circular and radial motions. In an early study with RDK stimuli by Freeman and Harris (1992), detection thresholds were measured to quantify the minimum motion needed to detect a global motion pattern. They found that circular and radial motions yielded lower detection thresholds than did translational motion, indicating that the visual system is more sensitive to complex motion. Edwards and Badcock (1993) added a speed gradient into radial motion patterns and found higher motion sensitivity for centripetal patterns (simulating contracting optic flows during backward self-motion) than centrifugal (simulating expanding optic flow during forward self-motion) and frontoparallel (translational) motion. However, studies using second-order (texture-defined) motion represented by RDK stimuli (Bertone & Faubert, 2003) found that observers were more sensitive in perceiving translational motion than complex motion patterns (i.e., circular and radial motions). A similar pattern of results was also reported in a study with stimuli requiring segregation of motion structure (Ahlstrom & Borjesson, 1996).

Most of the previous studies of motion integration have used random-dot kinematogram (RDK) as the experimental stimulus. A potential problem with using dot stimuli to study motion integration is that the observer might track the trajectory of a few signal dots in the display to infer the global motion direction. To discouraging this local-tracking strategy, researchers use the

limited lifetime technique with dot stimuli. However, this method does not completely rule out the tracking contribution if dot lifetime is more than two frames, or if the displacement of signal dot movement is small, or if signal dot movements can be easily segmented from noise dot movements. Furthermore, an inevitable consequence of introducing limited lifetime to dot stimuli is that it interferes with temporal smoothness, which could ultimately affect observer's sensitivity in perceiving global motion. Given these possible drawbacks, RDK stimuli may not be the ideal stimulus type for studying motion integration.

An alternative motion stimulus is an array of sinusoidal gratings, or lines with different orientations (Amano, Edwards, Badcock, & Nishida, 2009b; Mingolla, Todd, & Norman, 1992). Due to the aperture problem, the local velocity of each motion element is ambiguous when viewed individually. However, a coherent motion percept can be formed if the visual system processes motion information globally. This multiple-aperture stimulus can preclude local tracking and maintain temporal smoothness (as the Gabor gratings drift continuously) so as to amplify the use of sensitivity as a good measure of spatial integration. In addition, the multiple-aperture stimulus provides an effective tool for controlling the amount of information available at each element location, making it possible to tease apart different pooling mechanisms. For example, using the multiple-aperture stimulus, Amano and colleagues (2009b) found that the human visual system employs different pooling strategies adaptively depending on the ambiguity in the local motion signals. However, all stimuli in their study were translational motion. It is therefore unclear whether and how the human visual system could perceive global complex motion patterns, such as circular and radial motion, when viewing these multiple-aperture stimuli.

To understand the mechanisms of motion integration, we employed the multiple-aperture stimulus to compare human sensitivity for three basic global motion patterns: translational, circular and radial motion. We also investigated the characteristics of the motion integration mechanism for each specific motion pattern and examined the contributions of several factors that influence human global motion sensitivity. In the present paper we report three experiments. Experiment 1 was designed to compare motion sensitivity for translational, circular and radial motion over a large range of spatial density of local motion signals. Experiment 2 aimed to study the change of human motion sensitivity as a function of speed for each global motion pattern. Finally, Experiment 3 examined the time course of motion integration mechanisms for specific motion patterns.

Multiple-aperture stimulus

We adopted the stimulus created by Amano and et al. (2009b), who defined two terms, "1D motion" and "2D motion", for the purpose of stimulus description. The 1D motion (also termed "component motion") of a Gabor element refers to the drifting velocity of the element (as it can only drift in one single dimension orthogonal to the grating orientation). The 2D motion refers to the underlying true motion of the object behind apertures. For example, in Figure 2.1, the dark blue solid arrow in each aperture indicates the 1D motion of the drifting grating, whereas the thick pink arrows indicate the 2D motion of the whole object. 1D motion corresponds to observations obtained from each element from the multiple-aperture stimulus, and 2D motion is what observers need to infer about the true motion flow. To generate a multiple-aperture stimulus, the 1D motion of each Gabor element was computed based on its orientation and pre-assigned 2D motion direction.

The multiple-aperture stimulus consisted of 728 drifting Gabor elements arranged in a circular pattern inscribed in a 31 by 31 grid (as shown in Figure 2.2). Each Gabor element was an oriented sinusoidal grating windowed by a stationary Gaussian function. The spatial frequency of each grating was 5.58 cycles/deg and the standard deviation of the Gaussian window was 0.08 degree. Each Gabor element subtended a visual angle of 0.40 deg. The distance between the centers of two neighboring elements was 0.40 deg. The stimulus was displayed within a circular area of diameter 12.15 deg. There was a small, blank circular region with diameter subtending 1.96 degrees at the center of the display window. A red fixation dot was located at the center of the display. Contrast for all Gabor elements were set at 0.4. Orientation of each Gabor element was randomly assigned in each trial.

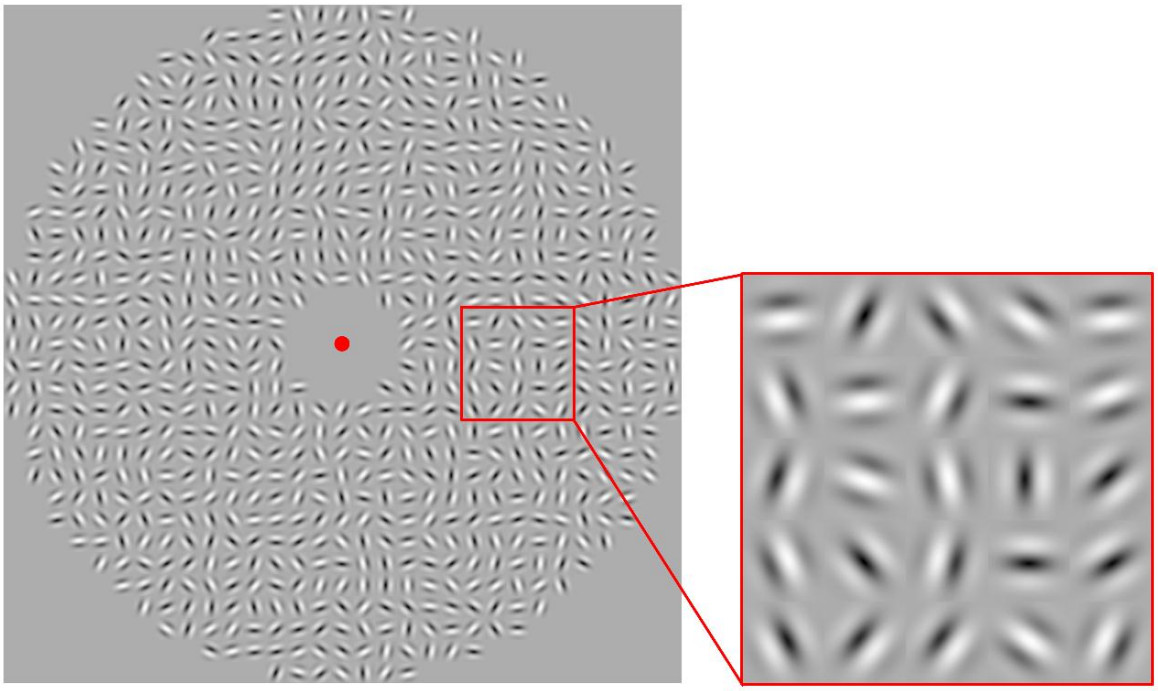

Figure 2.2. Screenshot of a typical stimulus instance and a zoomed view of the Gabor elements.

The three global motion patterns (translational, circular and radial motion) were created by manipulating 2D motion of Gabor elements. For translational motion, the 2D motion directions of all elements were the same (either rightward or leftward). For circular and radial motion, the 2D motion direction of each element was determined by its position relative to the center so that the pattern could be set to rotate around the center for circular motion (clockwise or counter-clockwise) or contract/expand relative to the center (inward or outward). Note that the speed of 2D motion was constant for all elements (implying non-rigid rotation and expansion/contraction

for complex motion), so as to be compatible with speeds in translational motion. Figure 2.3 illustrates the three global motion patterns used in this study.

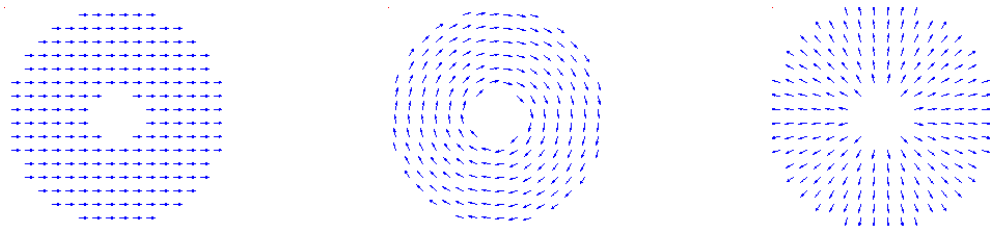

Figure 2.3. Schematic illustration of the three global motion patterns. Shown are three instances of motion patterns: rightward translational (left), clockwise circular (middle) and radially expanding motion (right). Arrows indicate the 2D motion vectors of the elements.

As shown in Figure 2.4, each Gabor element was categorized as either a signal or a noise element. For signal elements, the 2D motion was assigned as described in Figure 2.3, so that 2D motion for these signal elements represented a globally coherent motion pattern (translational, circular, or radial motion). For noise elements, the 2D motion directions were randomized while keeping the same speed. For both signal and noise elements (regardless of the motion pattern), the 2D motion speed was kept constant at 0.79 deg/s (except for Experiment 2).

Strength of global motion was controlled by coherence ratio, which was defined as the proportion of signal elements among the total number of elements in the stimulus. A stimulus with 100% coherence ratio would have all the elements as signals, whereas a stimulus with 50% coherence ratio would have half of the elements as signals while the other half were noise elements.

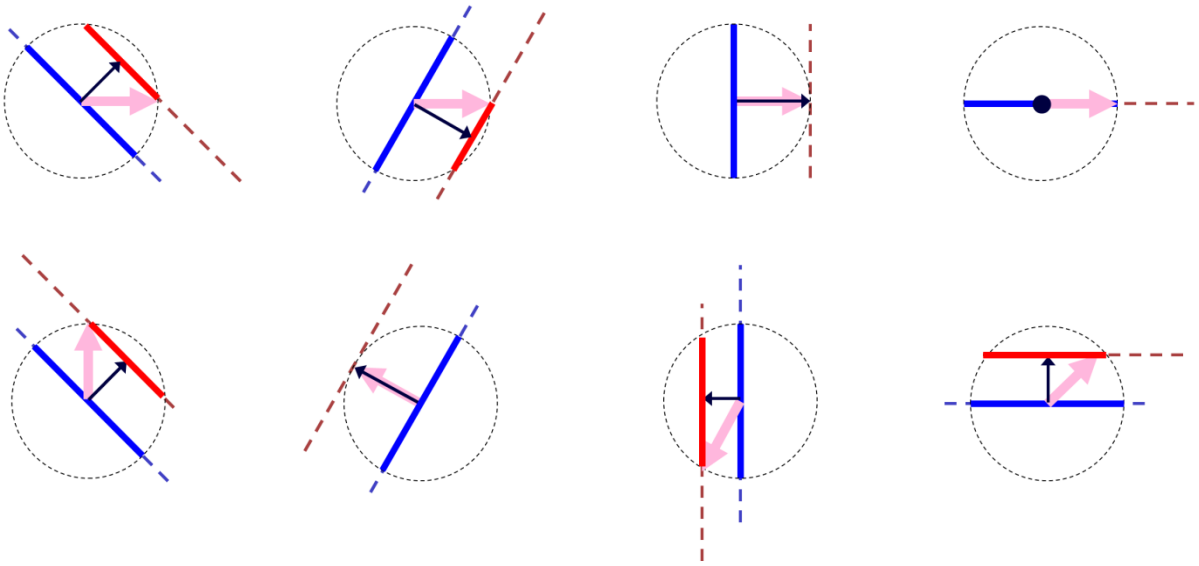

Figure 2.4. Diagrams of (upper) signal elements and (lower) noise elements. Thick pink arrows indicate 2D motion. Thin dark blue arrows indicate 1D motion. Same convention as used in Figure 2.1.

Experiment 1: Comparing human sensitivity for translational, circular and radial motion

Experiment 1 aimed to compare motion sensitivity in perceiving translational, circular and radial global motion using the multiple-aperture motion stimulus. If specific motion integration mechanisms operated for different motion types, we would expect to find the variation of human sensitivity for different motion types. A second goal of this experiment was to examine the integration characteristics in terms of the range of spatial pooling. The spatial extent of motion integration has been studied extensively using random-dot kinematorgrams (RDK). Several studies using the RDK stimuli demonstrated that human sensitivity for perceiving global

translation remains nearly constant when dot density is doubled (Edwards & Badcock, 1994, 1995, 1996), as well as when dot density varies over a relatively large range (Barlow & Tripathy, 1997). This density invariance effect was also found in complex motion, i.e., circular (Morrone et. al., 1995) and radial motion (Badcock & Khuu, 2001) using the RDK stimuli. Barlow and Tripathy (1997) provided a theoretical explanation of these findings using the ideal observer approach, which employs a linear pooling strategy within the display area to overcome the uncertainty induced by the correspondence problem inherent in the RDK stimuli. If a linear pooling strategy is a general integration mechanism adopted by the human visual system, we would expect the density invariance effect would be obtained using distinctly different motion stimuli. Accordingly, in Experiment 1 we measured motion sensitivity as a function of element density using the multiple-aperture stimulus, setting the size of the visual field to a value comparable to those used in previous studies with RDK stimuli.

Methods

Subjects. Fifteen undergraduate students from the University of California, Los Angeles (UCLA) participated in the experiment for course credit. The observers were naïve to the purpose of the experiment. All observers had normal or corrected-to-normal visual acuity. Five observers were randomly assigned to one of the three conditions: translational, circular and radial motion.

Apparatus. Motion stimuli were presented on a Viewsonic CRT monitor with a refresh rate of 75 Hz and resolution of 1024 × 768 pixels, with a constant viewing distance of 57 cm using a chin rest. Each pixel on the screen subtended 2.01 min-arc. A Minolta CS-100 photometer was used to calibrate the monitor. A luminance range of 0~146.5 cd/m^2 was converted into a linear

lookup table for 256 programmable intensity levels. Experiments were run in a dim room. Matlab and PsychToolbox (Brainard, 1997; Pelli, 1997) were used to present the stimuli.

Stimulus. The stimuli were similar to those described in the section 2, except for the spatial density. Different numbers of elements were presented at different levels of spatial density. We used seven density levels: 0.64, 1.29, 1.93, 2.58, 3.22, 4.84 and 6.45 elements/deg^2 (corresponding to 10%, 20%, 30%, 40%, 50%, 75% and 100% of 728 elements, respectively, shown in the display window). As depicted in Figure 2.5, Gabor elements were confined in a circular window but the positions of presented elements were randomly selected from trial to trial on the fixed grid.

On each trial, the motion sequence consisted of 20 frames with frame duration of 13.33ms per frame. Observers were told to fixate the red dot throughout the experiment. The fixation spot was presented for 500ms, and followed with the motion stimulus for 267ms. After the motion stimulus disappeared from the screen, observers were asked to press one of two keys to respond. Observers' task was to identify the global direction from two alternative directions for each motion type: left/right for translational, clockwise/counter-clockwise for circular, and inward/outward for radial motion. The key-pressing response triggered the start of the next trial after 1 second.

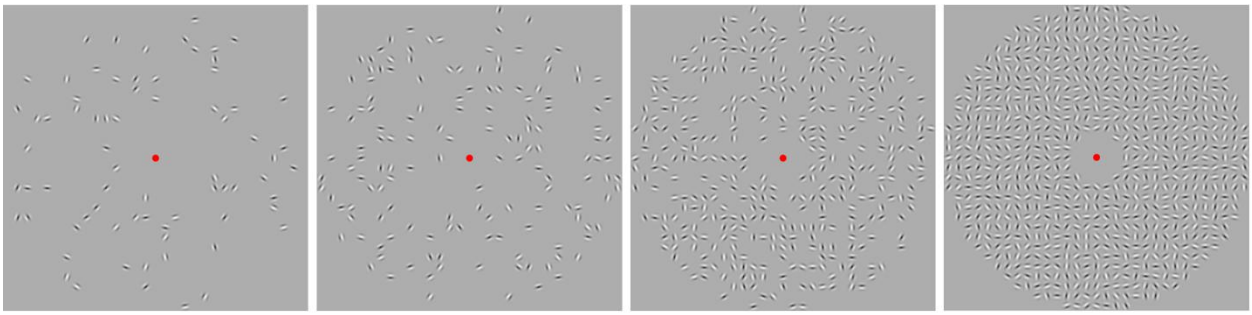

Figure 2.5. Stimuli of different densities used in Experiment 1. Examples illustrate four fixed levels of density of elements: (from left to right) 0.64 (10%), 1.29 (20%), 3.22 (50%) and 6.45 (100%) elements/deg^2.

Procedure. Seventy practice trials were administered to observers prior to each experiment session in order to familiarize them with the stimulus and the task. Coherence ratio was set to decrease during the practice, from 100% (first fourteen trials) to 80% (next seven trials), and then to 40% (the remaining 49 trials), so that observers could experience the different levels of task difficulty used in this experiment. After the practice, one experiment session was conducted, which included only one of the three types of motion (translational, circular or radial). In both practice and experiment sessions, the presentation order of the seven different density levels was randomized over trials. Within an experiment session, coherence ratio for each density level was independently adjusted to achieve 75% accuracy using the QUEST adaptive-staircase procedure (Watson & Pelli, 1983), so that coherence thresholds for the seven density levels could be estimated independently (150 trials for each level) and served as the motion sensitivity measurement. In every trial of both practice and experiment sessions, a beep was played to provide negative feedback whenever participants made an error.

Results and Discussion

As shown in Figure 2.6, the coherence threshold for translational motion was higher than that for circular/radial motion. High coherence thresholds indicate low sensitivity (i.e., worse performance in global motion perception). This result was confirmed by an analysis of covariance (ANCOVA) with motion types as a between-subjects factor, and density levels as a covariate. The main effect of motion types was found to be significant ($F(2, 101) = 11.94$, $p < .001$), meaning that the average thresholds for different motion types were found to be significantly different after adjustment for density levels. In particular, the coherence threshold for translational motion (adj. mean = .646) was found to be significantly different from those for circular (adj. mean = .437, $F(1, 101) = 18.68$, $p < .001$) and radial motion (adj. mean = .446, $F(1, 101) = 17.10$, $p < .001$), while that for circular motion was not significantly different from that for radial motion ($F(1, 101) = .035$, $p = .85$). This result reveals a "complexity advantage" for circular/radial motion over translational motion in terms of human sensitivity in perceiving global motion[1].

In assessing how spatial density affected motion sensitivity, we found that threshold-density slopes were not significantly different across motion types ($F(2, 99) = .036$, $p = .97$). Furthermore, the ANCOVA analysis revealed a non-significant main effect of element density ($F(1, 101) = .061$, $p = .81$), indicating that the density slopes for all motion types were not significantly different from zero. This result implies that motion sensitivity varied very little with

[1] This "complexity advantage" was found to be robust over a range of contrasts. In our pilot study, we employed 100% density and four levels of contrasts, 0.05, 0.1, 0.2 and 0.4 to compare human sensitivity for the three motion types. All other stimulus parameters were the same as in Experiment 1. The "complexity advantage" of circular and radial motion over translational motion was consistently observed across the four contrast levels for all three subjects, with only one exception when contrast was 0.05 for subject 3.

element density. The linear regression lines had mean slopes with 95% confidence interval as: –0.0045 ± 0.04 for translational motion, 0.0014 ± 0.02 for circular motion, and –0.0046 ± 0.05 for radial motion.

The results of Experiment 1 thus revealed lower motion sensitivity for translational than that for circular and radial motion. This finding suggests that specific motion integration mechanisms may operate for different motion types. Of particular interest is the counterintuitive observation that humans actually perform better in perceiving global circular/radial motion, motion types that have been considered more complex relative to translational motion. The results of Experiment 1 are consistent with those found by Freeman and Harris (1992) using the RDK stimulus. In terms of the effect of spatial density, the present findings are consistent with the results obtained using the RDK stimulus (Badcock & Khuu, 2001; Barlow & Tripathy, 1997; Edwards & Badcock, 1994, 1995, 1996; Morrone, et al., 1995), as human observers yielded similar qualitative patterns of results when using the multiple-aperture stimuli. This agreement indicates that the visual system is able to pool motion information over broad ranges with constant efficiency in perceiving translational, circular and radial motion. Nevertheless, the difference between simple translational motion and complex circular/radial motion remained across different density levels. In summary, the two findings, sensitivity difference across motion types and invariant sensitivity with element density, suggest that specialized integration mechanisms for the three motion patterns may share the same linear pooling principle as part of their computations.

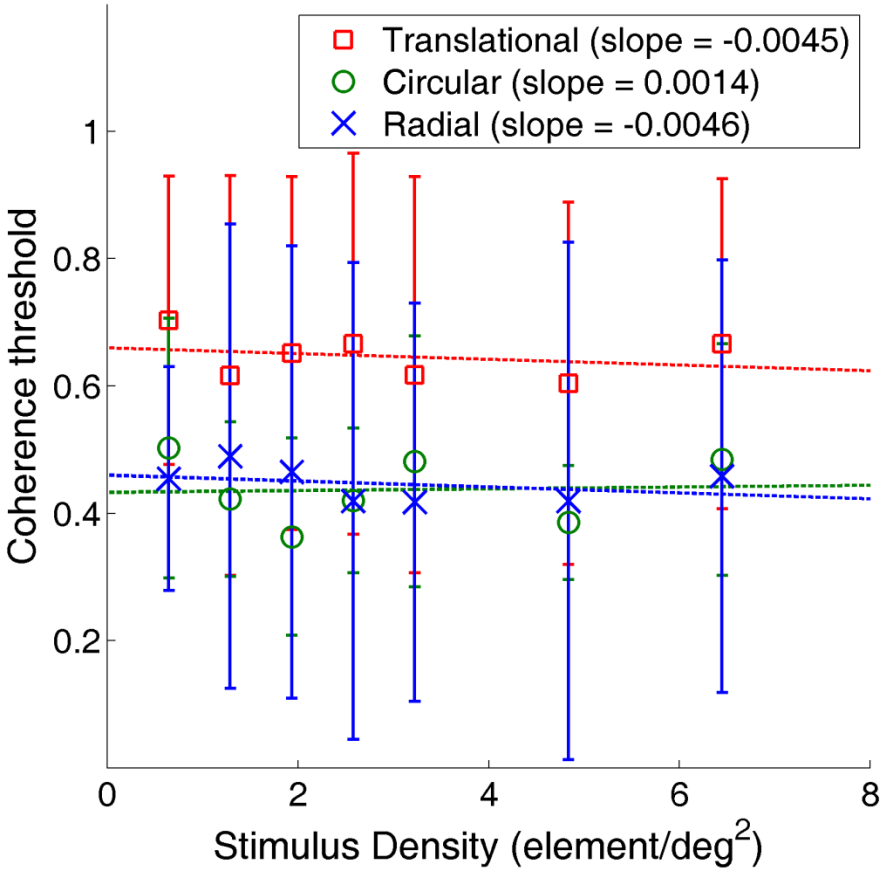

Figure 2.6. Results of Experiment 1. Coherence thresholds as a function of stimulus density for three motion patterns. Coherence thresholds (averaged across observers) are plotted across difference density levels for translational (red open squares), circular (green open circles) and radial (blue crosses) motion. For each motion type, the dotted line was fitted with averaged thresholds. Error bars show the 95% confidence intervals for the estimates of the means.

Experiment 2: Effects of speed on motion sensitivity

To investigate the dependency of sensitivity on motion speed for the three types of global motion, we measured motion sensitivity at various speeds using the multiple-aperture stimulus. Early physiological studies have demonstrated the speed tuning of motion-sensitive neurons. Maunsell and Van Essen (1983) found that MT cells in macaque monkeys responded to specific ranges of speed. Tanaka and Saito (1989) found that MST cells sensitive to translational and radial motion showed speed selectivity with a narrow speed tuning function. However, about one third of circular sensitive MST cells showed the property of speed invariance with a flat speed tuning function. Consistent results were obtained in subsequent studies on cells in MST within a similar speed range (Orban, Lagae, Raiguel, Xiao, & Maes, 1995), particularly in the dorsal region MSTd (Duffy & Wurtz, 1997). Experiment 2 aimed to investigate the effect of speed on motion sensitivity within the range of 0.7~ 2.5 deg/s. Given the fixed stimulus size and spatial frequency used in Gabor elements in the present study, faster speed tends to create a percept of flicker and this generates an unstable motion percept. We therefore focused on a low speed range relative to those tested in previous studies (Burr, Morrone, & Vaina, 1998; Orban, et al., 1995; Tanaka & Saito, 1989).

Methods

Subjects. A total of 102 undergraduate students at UCLA participated in this experiment for course credit. All observers were naïve to the purpose of the experiment, and had normal or corrected-to-normal visual acuity. Thirty-four participants were randomly assigned to one of three speed conditions.

Stimulus. The stimuli used in Experiment 2 were similar to those used in Experiment 1 with the highest spatial density, i.e., 728 elements. However, the 2D motion speed was varied for each observer group. Three speed levels were used: 0.79deg/s (same as that used in Experiment 1), 1.58deg/s, and 2.37deg/s. Speed was manipulated by changing the magnitude of phase shift of the Gabors per frame, so that the number of frames (and thus stimulus duration) remained the same as in Experiment 1.

Procedure. Observers performed the same global direction discrimination task as in Experiment 1. Each observer viewed motion stimuli for all three motion types but at one speed level. Observers first received a 60-trial practice session (20 trials for each motion type) to familiarize themselves with the stimulus and the task. Coherence ratio was decreased gradually from 100% to 90% over the 20 trials for each motion type. Participants were then presented with the experiment session, in which the three motion types were arranged in three blocks of 240 trials. The order of presentation of motion types was counterbalanced among observers. Each block started with on-screen instructions to inform observers of the motion type and response buttons in the upcoming block. In each block, participants were informed that the first 60 trials were practice for that particular type of motion. The remaining 180 trials were experiment trials, which were used to estimate coherence threshold via the QUEST adaptive-staircase procedure (Watson & Pelli, 1983). Participants were given negative feedback (a beep for error) in both practice and experiment trials, as in Experiment 1.

Results and Discussion

Figure 2.7 shows that the coherence threshold for translational motion was higher than those for circular and radial motion when speed was the slowest (translational vs. circular, $F(1, 99) = 5.25$,

p = .02; translational vs. radial, $F(1, 99) = 8.35$, p = .01), replicating the findings in Experiment 1. However, the difference between translational and radial motion was reduced as speed increased, although the thresholds for circular motion were nearly constant over the speed range tested in the experiment. A repeated-measures ANOVA with 2 factors (motion type as a within-subjects factor, speed as a between-subjects factor) was conducted to compare thresholds for different motion types at each speed level. This analysis revealed a significant main effect of motion type ($F(2, 198) = 23.88$, $p < .001$), indicating that the average sensitivity for the three motion types differed, confirming the general findings of Experiment 1. A significant interaction between motion type and speed was obtained ($F(4, 198) = 3.53$, p = .01), indicating that speed was an important factor influencing the sensitivity differences among motion types. Planned comparisons revealed how speed affects motion sensitivity for each specific motion type. In particular, global motion sensitivity for circular motion remained fairly constant across different speed levels ($F(2, 99) = .17$, p = .85), whereas sensitivity for radial motion was impaired as speed increased ($F(2, 99) = 4.80$, p = .01). A decreasing trend in sensitivity was also observed for translational motion as speed increased, although this trend fell short of statistical significance ($F(2, 99) = 1.39$, p = .26). These findings suggest that the integration mechanism specialized for circular motion may be tuned for a broad range of speeds. In contrast, radial motion integration mechanism may be more speed dependent. This implication is consistent with the physiological finding that a subgroup of circular MST cells (about one third of tested neurons) shows the speed-invariance characteristic, whereas most radial MST cells are speed-selective (Tanaka & Saito, 1989).

Within the tested speed range, the complexity advantage effect for circular versus translational motion demonstrated robustness over the change of the speed. However, the effect for radial

motion was speed dependent. From an ecological perspective, in natural scenes, radial motion is often associated with the observer's movements (ego-motion) under conditions rather different from those associated with translational and radial motion. In order to use the perceived optic flow to infer the observer's movement relative to the environment, the mechanisms for encoding radial motion needs to be rather sensitive to speed change. As a result, speed change may affect human sensitivity for radial motion more than it does for circular motion. Tanaka and Saito (1989) showed that average responses of all circular-motion-selective cells increase as speed increases; a similar trend was also obtained for radial-motion-selective cells. However, the increase in slope for radial motion cells was much sharper than the increase in slope for circular motion cells, which is consistent with our finding of speed-dependent sensitivity for radial motion and speed-invariance sensitivity for circular motion.

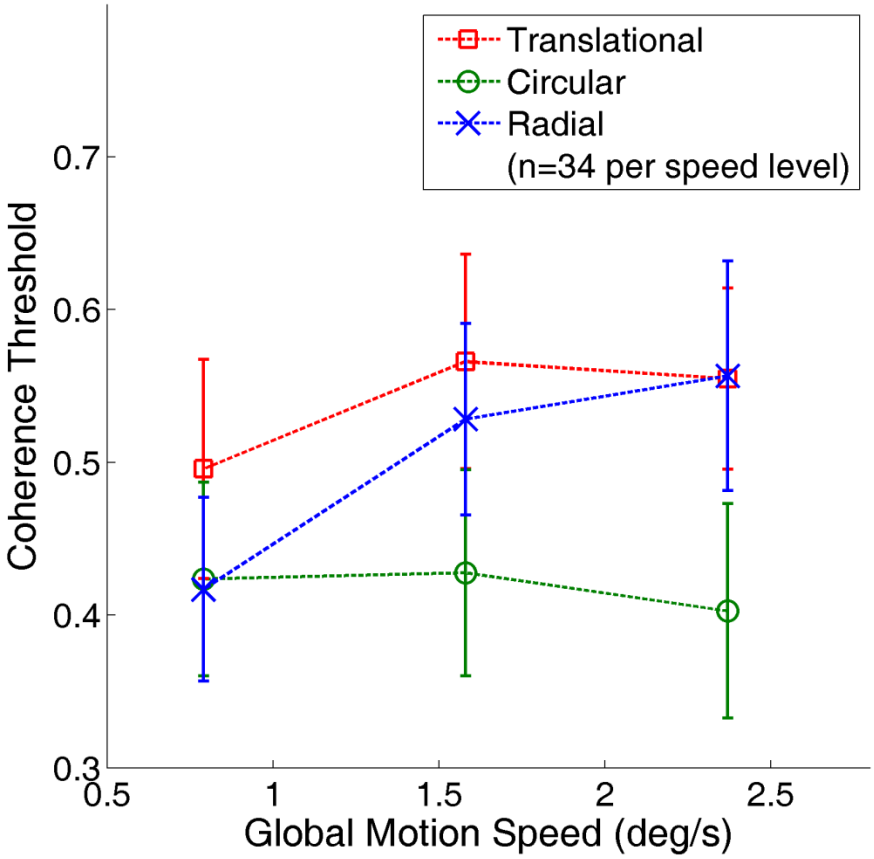

Figure 2.7. Results of Experiment 2. Coherence thresholds as a function of speeds. The figure shows the coherence thresholds (averaged across all observers) for translational (red open squares), circular (green open circles) and radial (blue crosses) motion. Error bars indicate the 95% confidence intervals for the estimates of the means.

Experiment 3: Time course of motion integration

Experiment 3 was designed to examine the temporal properties of motion integration mechanisms specialized for different motion types. Using random-dot stimuli, Burr and Santoro (2001) identified the "critical durations" needed for integrating different global motion patterns. Critical duration was defined as follows: if a motion stimulus is presented for a period of time that is shorter than the critical duration, motion sensitivity increases linearly with stimulus duration on a log-log scale; when the presentation time is longer than the critical duration, motion sensitivity remains constant when stimulus duration is further increased. Burr and Santoro (2001) used coherence threshold as the measure of motion sensitivity using the RDK stimulus. The investigators found that the global-motion integration stage needs long summation periods, about 3 seconds. In addition, their Experiment 2A revealed comparable critical durations for the three motion types. This result motivated us to conduct Experiment 3 using the multiple-aperture stimulus to examine the time course of integration for different motion types.

Methods

Subjects. Four observers (the first author and three naïve subjects) participated in this experiment. All observers had normal or corrected-to-normal visual acuity.

Stimulus. The stimuli were identical to those with the highest spatial density used in Experiment 1, except that the presentation time of the motion stimulus was manipulated. Fourteen different stimulus durations between 27 ms (two frames) and 413 ms (31 frames) were tested.

Procedure. Observers performed the same global direction discrimination task as in Experiment 1. Each observer viewed motion stimuli for all three motion types, with each motion type presented in a separate session. Observers were informed of which motion type they would be

viewing in a particular session. The order of motion types were counterbalanced between observers. Before the start of each experiment session, observers were given 112 practice trials of the same motion type, with randomized stimulus durations (eight trials for each of the 14 stimulus durations). Coherence ratio was 100% for all practice trials. During the experiment session, observers completed a total of 1400 trials, with the order for the 14 durations randomized from trial to trial. QUEST adaptive-staircase procedure was used to estimate coherence thresholds for the 14 stimulus durations (100 trials for each duration). Participants were given negative feedback in both practice and experiment trials as in Experiments 1 and 2.

Results and Discussion

Figure 2.8 depicts the results from four observers in Experiment 3. We performed an analysis similar to that conducted by Burr and Santoro (2001). Motion sensitivity was defined as the reciprocal of coherence threshold. Least-square regression was performed to estimate the parameters in the following equation for sensitivity (adopted from Burr & Santoro, 2001):

$$S = \begin{cases} S_0 \left(\dfrac{t}{t_0} \right)^p & \text{for } \; t \le t_0 \\ S_0 & \text{otherwise} \end{cases},$$

where S indicates measured motion sensitivity and t indicates duration of motion stimulus. The parameters included p, the slope of the fitted line in a log-log coordinates; t_0, the critical duration for temporal integration; and S_0, a constant for motion sensitivity at the plateau level of performance. The three parameters were estimated using nonlinear regression, with a constraint that the predicted sensitivity was greater than or equal to one (because the coherence ratio is bounded within a range of 1 to 0) for the shortest duration (i.e., 27 ms) in the experiment.

The three parameters were estimated for each motion type for each observer. Data points with standardized residual greater than or equal to 2 in the regression analysis were categorized as outliers. Based on this criterion, one data point was excluded in the regression analyses in the circular condition for observer AL, in the translational condition for observer OQ, and in all three motion conditions for observer NH. As shown in Table 2.1. Estimated parameters in regression analysis to describe human sensitivity as a function of stimulus duration.

The parameters *are* p, the slope of the fitted line in a log-log *coordinates;* t0, the critical duration for temporal integration, *and* S0, a constant for motion **sensitivity.**, the overall goodness of the fit, R^2 for the regressions was in the range from 0.71 to 0.92, with an average of 0.85. These findings indicate a reasonable quality of fit.

Obs.	Motion type	p	95% C.I. for p	t_0	95% C.I. for t_0	S_0	95% C.I. for S_0	R^2
AL	**Translational**	0.73	(0.24, 1.22)	131.76	(71.63, 191.88)	3.20	(2.86, 3.54)	0.82
	Circular	1.12	(0.47, 1.76)	93.01	(67.31, 118.71)	4.04	(3.72, 4.37)	0.90
	Radial	0.85	(0.29, 1.40)	115.08	(74.33, 155.82)	3.46	(3.05, 3.88)	0.79
JC	**Translational**	0.52	(0.20, 0.84)	156.00	(77.75, 234.25)	2.51	(2.25, 2.78)	0.81
	Circular	0.78	(0.41, 1.16)	160.00	(108.74, 211.26)	4.08	(3.55, 4.60)	0.86
	Radial	0.82	(0.45, 1.19)	161.83	(113.26, 210.40)	4.38	(3.84, 4.92)	0.88
NH	**Translational**	0.16	(0.03, 0.29)	106.67	(36.73, 176.61)	1.25	(1.19, 1.31)	0.71
	Circular	1.31	(0.48, 2.13)	84.78	(64.49, 105.07)	4.54	(4.08, 5.00)	0.87
	Radial	0.45	(0.26, 0.64)	165.42	(101.39, 229.44)	2.25	(2.03, 2.48)	0.88
OQ	**Translational**	0.69	(0.39, 0.98)	207.98	(130.17, 285.79)	4.09	(3.55, 4.64)	0.90
	Circular	1.35	(0.66, 2.04)	154.69	(113.04, 196.33)	10.70	(9.57, 11.84)	0.92
	Radial	1.03	(0.43, 1.64)	130.43	(86.36, 174.51)	5.15	(4.64, 5.66)	0.90

Table 2.1. Estimated parameters in regression analysis to describe human sensitivity as a function of stimulus duration. The parameters are p, the slope of the fitted line in a log-log coordinates; t_0, the critical duration for temporal integration, and S_0, a constant for motion sensitivity.

Table 2.1 shows the parameter estimates for each motion condition for each observer. We found that the estimated critical durations for temporal integration (t_0) for all motion types were short, around 140 milliseconds. This result is not in agreement with the temporal limits reported by

Burr and Santoro (2001), in which they found the critical durations for all motion types to be about 2~3 seconds (more than 10 times longer than the critical durations as revealed in the present study). This large discrepancy may be due to the use of different stimuli (multiple-aperture stimulus versus RDK stimulus) and experimental setups, a point to be taken up in the discussion. On the other hand, we found that critical durations were not significantly different for different motion types; the 95% confidence intervals of estimated critical durations for the three motion types overlapped for all observers, as shown in Figure 2.8. This result (invariance in the limit of temporal integration across motion types) was consistent with findings from the RDK experiment (Experiment 2A) reported by Burr and Santoro (2001). For the saturated range of duration (longer than critical duration), motion sensitivity for complex motion types (i.e., circular and radial) was found to be higher than that for translational motion, which again replicated the main findings from Experiments 1 and 2.

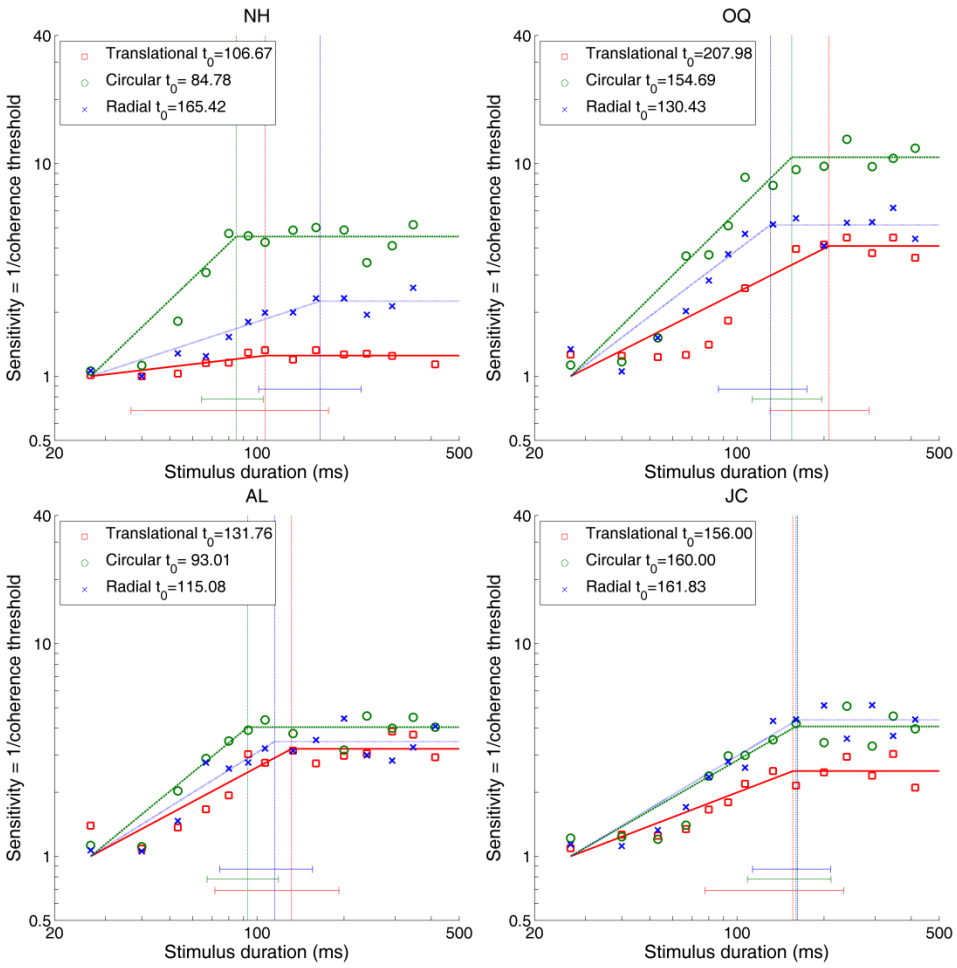

Figure 2.8. Results of Experiment 3. Motion sensitivity for each motion type (separate lines) as a function of stimulus duration for each observer (separate graphs). Motion sensitivities for translational, circular and radial motion are plotted as red open squares, green open circles and blue crosses respectively. Each fitted line (translational: red solid; circular: blue dash-dot; radial: green dashed) shows the predicted sensitivity for each motion type using the fitted parameters. Vertical dashed lines mark the critical duration (estimate of t_0) for each motion type, and the horizontal error bars mark the 95% confidence intervals of the estimates in their corresponding colors.

General Discussion

The results of all three experiments reported in the present study demonstrate greater motion sensitivity in perceiving complex motion (i.e., circular and radial motion) than in perceiving translational motion, implying that different motion integration mechanisms are involved in the computation for different global motion patterns. This general finding is consistent with that reported by Freeman and Harris (1992), based on a study with the random dot stimuli. Despite the difference between multiple-aperture stimuli used in the present study and the dot stimuli used by Freeman and Harris (1992), both studies were similar in using motion stimuli at relatively slow speed, which might be the reason that this complexity advantage was obtained in both studies. Furthermore, the qualitative patterns of the present results were also observed in studies using Glass patterns to investigate global form perception (Wilson & Wilkinson, 1998), in which observers were found to be more sensitive in detecting concentric and radial patterns than parallel patterns. As pointed out by Ross, Badcock and Hayes (2000), form perception of Glass patterns might be closely related to motion perception of corresponding motion types at higher level of visual processing. Taken together with these studies, our findings provide converging evidence to support a complexity advantage in perceiving different types of motion patterns.

In contrast, Blake and Aiba (1998) did not find a significant difference in motion sensitivity for translational, circular and radial motion. A possible account for this discrepancy could be the difference between the stimuli used in the two studies. Given their goal of studying general optic flow processing, Blake and Aiba (1998) used densely distributed dots moving against a bright background as motion stimuli, which produced correspondence noise (H. Barlow & S. P. Tripathy, 1997). From a computational perspective, correspondence noise introduces labeling

uncertainty in determining which dot in one frame corresponds to which dot in the next frame (H. Barlow & S. P. Tripathy, 1997; H. Lu & Yuille, 2006). As a result, correspondence uncertainty directly affects the local velocity estimate of each dot and further influences the perceived global motion via spatial integration. However, given that 1D motion (component motion) is observable for the multiple-aperture stimulus used in the present study, the visual system needs to adopt an appropriate pooling strategy to integrate the observed 1D motion over space in order to infer the global motion. The inconsistency in findings between the present study and that of Blake and Aiba (1998) may thus reflect the fact that different stimuli produced uncertainty at different levels of motion processing.

Another discrepancy involving human performance when using the RDK stimuli compared to the multiple-aperture stimuli was revealed in Experiment 3. Whereas observers need 2~3 seconds for temporal integration to perceive global motion displayed in the RDK stimulus (Burr & Santota, 2001), the multiple-aperture stimulus yielded a much shorter period required for temporal integration, 100~200 milliseconds. We suspect that different levels of uncertainty introduced by the two distinct stimuli could account for this large discrepancy. As each element in the multiple-aperture stimulus was a continuously drifting Gabor grating with medium contrast, local motion processors (such as motion energy detectors) could readily extract the component motion velocity (1D motion) for each element. In contrast, the RDK stimulus introduces large correspondence uncertainty in the estimation of the local velocity of each dot. Several studies have shown that the human visual system is not efficient in coping with this type of noise, as indicated by very low absolute efficiency value, less than 5% (Barlow & Tripathy, 1997; Lu & Yuille, 2006). Thus the visual system needs more stimulus information, as is provided by longer duration, to perceive global motion displayed in the RDK stimulus.

Although humans showed shorter durations for temporal integration using the multiple-aperture stimulus than using the RDK stimulus, the absolute coherence thresholds (.4~.6) obtained in our study was much higher than the thresholds (.05~.15) that have been reported in most studies using the RDK stimulus. This discrepancy may be due to the use of local-tracking strategies, which may facilitate global motion perception for RDK stimuli in some studies. Also, in most previous studies observers were well-trained and experienced participants in psychophysical experiments. The observers in our studies, junior undergraduate students who participated for course credit, were inexperienced and likely less motivated. Although we provided 70 practice trials before the experiment, naïve observers still produced much higher thresholds than experienced ones. In our pilot studies using experienced observers (n=7), average discrimination thresholds for translational, circular and radial motion were 0.35, 0.20 and 0.27, respectively. These values are about half of the thresholds obtained for naïve subjects. These results have been reported in abstract form (Lee, Yuille & Lu, VSS, 2008). The differences in terms of subject population also contributed to the finding that thresholds in our studies were much higher than those found in previous studies by Amano et al. (2009) using the multiple-aperture stimulus in a large viewing window, which were in the range of .1~.3. The fact that the complexity advantage is obtained even with naïve subjects indicates the robustness of the phenomenon.

In addition, three factors related to stimulus parameters may also have influenced the main findings in our study. First, eye movements may interfere with translational motion more than the other two complex motions, which might contribute to the complexity advantage found in our study. Although we cannot completely rule out this possibility, we believe the effect of eye movements was relatively small. Most of our stimuli had a brief presentation duration, around 260 ms (except the shortest was 27ms and the longest was 413 ms in Experiment 3), minimizing

the effect of eye movements. Also, all observers were given clear instructions to fixate the central dot throughout the session, and had sufficient practice trials before they ran any experiment sessions.

A second possible concern involves the removal of speed gradient for circular and radial stimuli. Lack of speed gradient for the two complex motions leads to a percept of non-rigid rotation and expansion, which is inconsistent with optic flows observed in natural viewing situations. Although this is a very reasonable concern, this lack of naturalness may have placed the complex motion stimuli at a disadvantage for sensitivity, and hence cannot explain the complexity advantage we obtained. In addition, we conducted a pilot study (reported in abstract form, Lee, Yuille & Lu, 2008) that compared human sensitivity when viewing rigid (with speed gradient) and non-rigid (without speed gradient) movements for circular and radial motion using multiple-aperture stimuli[2]. We did not find any significant difference in motion sensitivity between rigid and non-rigid patterns. This finding is consistent with evidence from early physiological studies (e.g., Tanaka, Fukuda & Saito, 1989), which found that speed gradient did not significantly affect responses of selected neurons in the MST area of macaque monkeys (neural responses were similar for motion patterns with and without speed gradient). Therefore, we chose to keep the 2D speed constant for all motion types so as to make our findings comparable with translation, and with previous studies using the similar method (e.g., Burr & Santoro, 2001; Morrone, Burr & Vaina, 1995).

[2] Ten naïve subjects participated in this pilot study. Each observer participated in five blocks, including translational motion, circular motion with and without speed gradient , and radial motion with and without speed gradient. The other stimulus parameters and experimental procedures were the same as in Experiment 1, using the largest density level. We found no difference between rigid and non-rigid conditions. Thresholds were respectively .46 and .42 for the two conditions with circular motion, and .45 and .42 respectively with radial motion.

A third possible explanation for the complexity advantage is that the difference in motion sensitivity between translational and radial motion might be related to the inward bias found by Edwards and Badcock (1993). Because the present study employed the two-alternative direction discrimination task, the higher sensitivity of radial motion might be due to the higher sensitivity of inward motion on half of the radial trials, with translational and the other outward half of radial motion having similar sensitivity. However, if this were the explanation, then such a difference in motion sensitivity between translational and radial motion would have been observed in previous studies using a similar task. In fact, findings from previous studies using a similar procedure (e.g., Aaen-Stockdale et al., 2007; Bertone & Faubert, 2003; Morrone et al., 1995) reveal no significant difference between translational and radial motion, suggesting that the inward bias does not account for the general effect. In summary, it does not seem likely that the complexity advantage observed in the present study can be largely attributed to any of the above-mentioned experimental or procedural factors.

The key finding of the present study, the complexity advantage, may seem counterintuitive in that the visual system was found to perform better on a more complex task (circular or radial motion cannot be defined simply by a unidirectional motion flow across space, as is the case for translational motion). However, physiological studies suggest that several high level visual areas are involved in processing complex motion patterns, which may help increase motion sensitivity for circular/radial motion in behavioral tasks. Consistent with the ecological importance of perceiving circular and radial motion, early physiological studies (Duffy & Wurtz, 1991a; Tanaka & Saito, 1989) found that some neurons in the MST region selectively respond to circular/radial patterns. Furthermore, other studies have identified neurons that respond specifically to rotational and expansion patterns in other brain areas responsible for high level

processing. Sakata et al. (1994) found that some "rotation-sensitive" neurons in the posterior parietal (PP) area respond specifically to rotary movements of objects. Krekelberg et al. (2003) found that implied motion from dynamic Glass patterns (Ross, et al., 2000) triggered responses in neurons in the high-level motion area superior temporal sulcus (STS). Wall and Smith (2008) identified the cingulate sulcus visual area (CSv) as a candidate region that seems to respond exclusively to expanding flow patterns that imply egomotion (but not to other expanding patterns that do not imply egomotion). Evidence from these studies supports the hypothesis that there exist high-level processing units tuned for circular or radial motion, which may provide a top-down influence that aids selection of an appropriate integration mechanism tailored to specific motion patterns in order to perceive global motion.

In addition to the involvement of high level visual areas in analyzing circular/radial motion patterns, more neural resources may be allocated to processing complex motion as opposed to simple translational motion within the MST area, which has long been recognized as the primary area for processing optic flow. For example, Tanaka and Saito (1989) clustered MST cells into three types of responsive groups: direction cells that selectively respond to one of *eight* translational directions; rotation cells that selectively respond to *two* directions (either a clockwise or counterclockwise rotation); and expansion/contraction cells that selectively response to *two* directions (either an expansion or contraction). Although the total number of cells appeared to be similar across the three clusters, the number of cells specifically selective to left or right motion was less than the number of cells selective to one direction in circular and radial motion (since the researchers used *eight* translational directions for their translational stimuli, but only *two* directions for their circular and radial motion stimuli). The similar argument could also apply to the study (Duffy & Wurtz, 1991), which used *four* translational

directions and *two* directions for complex motion. If a population-encoding strategy were adopted for neuronal processing of optic flow analysis, the variance of averaged response from selective neurons would be reduced with an increase in the number of selective neurons. Accordingly, we conjecture that the involvement of more MST neurons selective to circular/radial motion might improve the signal-to-noise ratio in neural processing, which could increase behavioral sensitivity and thus lead to the complexity advantage reported in this paper. Our main finding, the complexity advantage, refutes any model in which the same rules for integrating local velocity signals are used in performing motion analysis for different global motion patterns. Several studies have shown that a generic integration strategy, such as a slow-and-smooth prior, can predict human performance well for translational motion (Weiss, Simoncelli, & Adelson, 2002; Yuille & Grzywacz, 1988). However, Wu, Lu and Yuille (2009) showed that the standard slow-and-smooth prior cannot be a universal integration strategy for other motion types. Instead, for circular and radial motions, registration of motion structure was required, which can be modeled by extending the smoothness term in the generic priors. In their model, the circular and radial smoothness priors do not include any information about the specific rotation/expansion center (analogous to the standard slow-and-smooth prior in which no specific global translational directions are encoded). Thus the model does not deliberately include any bias favoring complex motions over simple translational motion. Nonetheless, simulations confirm that this model is able to select the most effective integration strategy based upon perceived motion information, and this choice in turn affects the estimation of motion flow. This model further predicts the complexity advantage observed in our study, and provides a good correspondence with neural processors specialized for complex motion, as revealed in physiological studies.

In addition, due to the general preference for slow motion encoded by the slowness term in all three generic priors, Wu et al.'s (2009) model predicts a general trend of reduced sensitivity (i.e., higher coherence thresholds) with an increase of speed. This prediction is largely consistent with the finding that speed affected coherence thresholds, as shown in Experiment 2. The fact that the influence of speed varied for different motion types (as shown Figure 2.7) suggests that relative weights between slowness and smoothness in the three generic models may differ depending on specific motion types. Future computational studies are needed to verify this conjecture by comparing with natural statistics of translational, circular, and radial optic flow fields.

Chapter 3

Multilevel Motion Adaptation

Introduction

Previous neurophysiological studies have found that, after prolonged stimulation, neurons in the visual cortex change their response characteristics (Dragoi, et al., 2000; Kohn & Movshon, 2003, 2004; Krekelberg, et al., 2006), and that such adaptation-induced neural changes are related to remarkable perceptual aftereffects (for a review: C. W. Clifford, et al., 2007; Schrater & Simoncelli, 1998; e.g., F. A. Verstraten, et al., 1994). The special case of contrast adaptation has been extensively studied over the past 20 years. Converging physiological and psychophysical evidence has shown that contrast adaptation induces neural changes through two mechanisms. First, adaptation independently modulates neural activity at early processing stages (retina and V1; (Baccus & Meister, 2004; Solomon, et al., 2004); second, these low-level changes are propagated up the visual hierarchy to affect neural responses in higher-level areas such as MT (Kohn & Movshon, 2003). It remains unclear whether these two basic mechanisms are also recruited for other types of sensory adaptation, and whether adaptation-induced neural changes at different processing levels can lead to distinct perceptual aftereffects. To address these questions, the present study investigates a different form of adaptation: motion adaptation. We used transparent motion as an adapting stimulus to trace the effects of adaptation through the motion processing hierarchy.

A transparent motion stimulus (Qian & Andersen, 1994; F. A. Verstraten, et al., 1994) contains multiple motion components[3] overlapping in the same spatial region. There is evidence that the processing of transparent motion involves multiple levels of motion analysis, including the extraction of local motion signals based on detectors with small receptive fields, such as V1 neurons, followed by a pooling of these local measurements through spatial integration based on detectors with large receptive fields as MT neurons (Qian & Andersen, 1994; Qian, et al., 1994a, 1994b; Snowden, Treue, Erickson, & Andersen, 1991). Transparent motion, thus, provides a useful tool to investigate how adaptation affects multiple stages within the motion processing hierarchy.

In addition, there is a puzzling perceptual effect in transparent motion adaptation. Although observers can simultaneously perceive multiple component motion directions when adapting to transparent motion (Snowden & Verstraten, 1999), such adaptation generally elicits a unidirectional, integrated motion aftereffect (MAE) opposite to the average of the adapting directions (Mather, 1980; Snowden & Verstraten, 1999; F. A. Verstraten, et al., 1994). This integrated MAE percept after adapting to bidirectional transparent pattern was first explained by the distribution-shift model (Mather, 1980). However, this model cannot explain why multiple directions in the transparent stimulus can be readily segmented and simultaneously perceived (Grunewald & Lankheet, 1996). Therefore, it remains unclear why motion segmentation is performed during transparent motion adaptation, yet such segmentation information does not influence the subsequent aftereffect to produce transparent MAE (Snowden & Verstraten, 1999).

[3] The word "component" may lead readers to think of the "component vs. pattern" terminology used in the motion processing literature. To avoid confusion, the word "component" in this paper is used solely for referring to one of the multiple motion directions embedded in a transparent motion display, and does *not* refer to the local drifting motion of gratings.

Two competing theories have offered different explanations for why adaptation to multiple directions of motion yields a unidirectional, integrated MAE under most conditions. Grunewald and Lankheet (1996) proposed that transparent motion adaptation modulates the interactions between neurons tuned to different global motion directions. In this view, adaptation-induced changes primarily occur at the stage of global motion processing through a broadly-tuned, inhibitory mechanism. In contrast, Vidnyanszky and colleagues (2002) proposed that prolonged exposure to a transparent motion stimulus induces *bidirectional* local motion signals at each position and causes local mechanisms to average the different motion signals, resulting in unidirectional local aftereffects at each position. At the higher-level, motion-integration stage, these local aftereffects are integrated over space to generate the unidirectional MAEs. This theory, thus, supports the hypothesis that transparent motion adaptation involves an essential step of local aftereffect integration. Consistent with this local processing account, Curran, Clifford and Benton (2006) used unidirectional random-dot kinematograms to show that the perceived directional aftereffect is mainly driven by the adaptation of motion-sensitive cells at the local-processing stage of motion analysis. In addition, a recent study by Scarfe and Johnston (2011) provided compelling evidence that a unidirectional moving pattern can bias the perceived local aftereffects, suggesting that low-level detectors not only project motion signals to, but also receive feedback from, cells involved in high-level motion processing.

The present study aims to reconcile these two accounts for transparent motion adaptation within a coherent framework. We developed a novel experimental paradigm to show that adaptation to bidirectional transparent motion patterns can lead to two radically

different types of MAEs: segregated and integrated MAEs. The segregated MAE yields an aftereffect opposite to one of the adapting directions in the transparent motion stimulus. In contrast, the integrated MAE yields an aftereffect opposite to the average of the adapting directions. Experiment 1 shows that when local-level adaptation is strong, local adaptation effects are propagated to higher-level motion processing. The integration of these local aftereffects gives rise to the percept of a segregated MAE, which possibly overrides integrated MAE resulting from adaptation-induced modulation at the global level. Experiments 2 and 3 show that, when local adaptation effects are eliminated or weakened, perceived MAE is dominated by the global-level adaptation effects, resulting in an integrated MAE.

Methods

Stimuli

We employed a multiple-aperture stimulus composed of multiple randomly-oriented, drifting sinusoidal gratings (Amano, Edwards, Badcock, & Nishida, 2009a; Clark & Bradley, 2008; Lee & Lu, 2010), as shown in Figure 3.1A. Grating elements drifted within fixed windows to generate local motion signals, which could produce strong local adaptation effects after prolonged viewing. These illusory motion signals of local aftereffects show different motion directions and speeds because grating orientation was randomly assigned for each location. Drifting speed of each element was controlled to generate globally coherent motion through spatial integration. A transparent motion display can be constructed by randomly grouping the elements into two sets (as color-coded in Figure 3.1) and assigning two different global motion directions to each of the two sets. The resulting percept resembles those described in previous

studies of transparent motion using random-dot kinematograms, in which two segregated motion directions can be perceived (Amano, et al., 2009a).

The multiple-aperture stimulus (Figure 3.1A) consisted of 396 drifting Gabor elements arranged in a circular pattern inscribed in a 24×24 grid, without any separation between each cell. Each Gabor element, subtending a visual angle of 1°, was constructed by imposing a stationary Gaussian function over an oriented sinusoidal grating, with spatial frequency being 2 cycles/° and the standard deviation of the Gaussian window being 0.3°. Distance between the centers of two adjacent Gabors was 1°. The stimulus was displayed within an annulus spanning 4°–12° around fixation. As population receptive field size of V1 in the human brain has been found to be under 1° within our stimulus range of eccentricity (Dumoulin & Wandell, 2008), integration of multiple elements by a local motion detector in V1, if any, should be minimal.

Orientation of each Gabor element was randomly assigned on each trial. For each Gabor element, the local drifting speed *u* was computed as follows:

$$u = v \sin(\alpha - \theta)$$

, where α is the global motion direction, *v* is the global motion speed, and θ is the orientation assigned to the element. This way of computing the local drifting speeds ensures that local motion vectors of all elements are consistent with the assigned global motion. Low contrast (0.05) for all Gabor elements was used to elicit strong spatial integration to facilitate the perception of global motion from the multiple-Gabor stimulus (Amano et al., 2009).

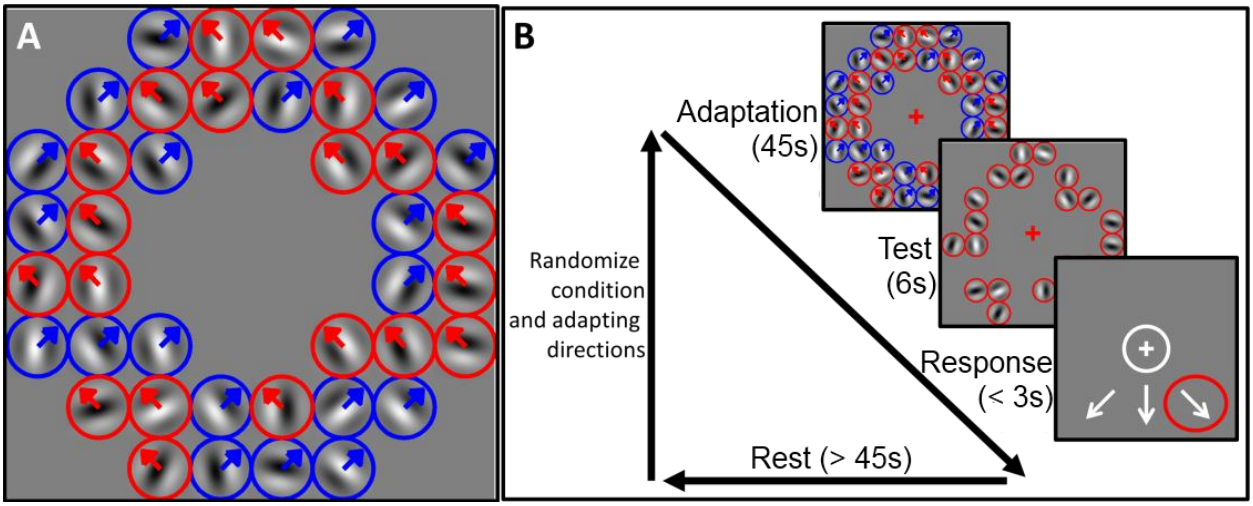

Figure 3.1. Illustrations of the stimulus and general procedure. A: The adapting stimulus consisted of multiple drifting grating elements. Colors code the two sets of elements to which elements are randomly assigned (Red: Set 1, Blue: Set 2). The arrow overlaid on each element indicates the assigned global motion vector (Up-Left for Set 1 and Up-Right for Set 2) B: Procedure for a sample trial (a *Single* condition). Colors and arrows on Gabor elements are for illustration only and were not presented during the experiments. Number of elements is reduced for the purpose of illustration.

The transparent motion pattern

All transparent motion patterns in the present study contained two component motion directions. Each Gabor element was randomly assigned to one of two sets, and each set was assigned with component motion vector. As a result, the two component sets of elements were spatially separated with random configuration. Elements in set 1 were assigned a global motion direction of (X-45)°, while elements in set 2 were assigned (X+45)°, in which X° indicate the averaged or integrated direction of two component velocities. Eight integrated motion directions

(from 0° to 315°, with 45° separation) were used in the experiments. Global motion speed was set at v = 3.15°/s for both component vectors.

This method of generating transparent motion pattern was adopted from the study by Amano et al. (2009). The main difference was that they employed a transparent motion stimulus with two component directions 180° apart (i.e., opposite), whereas ours were always 90° apart. Amano et al. showed that orientations of Gabor elements need to be randomized in order to perceive transparent motion. In a separate experiment, we verified that observers can reliably identify the two individual directions embedded in this stimulus. Eleven observers were instructed to indicate as many directions as they perceived on the transparent pattern. They indicated their perceived directions, one after another, by turning a simulated dial on the computer screen. Results are shown in Figure 3.2. In the majority of trials, observers indicated that they perceived two directions. The reported perceived directions were found to peak at two directions, which generally coincide with the two directions embedded in the adapting stimulus. This finding justifies the use of this stimulus as a bidirectional, transparent motion pattern.

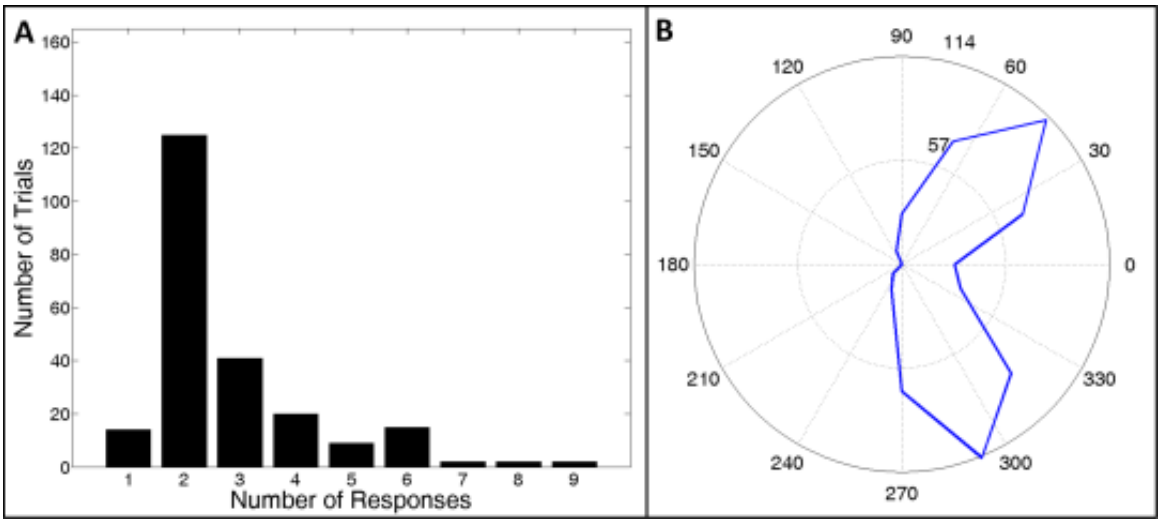

Figure 3.2. Distributions of perceived directions on the multiple-aperture transparent pattern in the dial-turning task. The stimulus contained two global directions that are 90° apart. A: Distribution of numbers of responses across all observers (230 trials in total). B: Distribution of perceived directions. Angles represent reported directions. Distance from center represents number of trials. Responses are aligned so that the two embedded directions in the stimulus are 45° and 315° (i.e., –45°).

General procedure

Each experimental trial consisted of three phases: adaptation (45 seconds), test (6 seconds) and response (< 3 seconds) (Figure 3.1B). First, during adaptation observers viewed the transparent motion pattern, with fixation maintained at the central cross. Then, they were presented with a static test pattern, with elements taken directly from the last frame of the adaptation motion sequence, while maintaining fixation at the central cross. Observers indicated their perceived motion direction of the aftereffect by choosing one of the following 4 options: 3

directional responses and 1 “no-motion” response, as illustrated in Figure 3.1B. Suppose a trial was assigned the integrated direction X°. Then, the adapting directions of component set 1 and set 2 were (X-45)° and (X+45)°, respectively. The three directional response options were always (X+135)°, (X-135)° and (X-180)°, which were opposite to set 1, set 2 and the integrated adapting directions, respectively. The first two responses were defined as segregated MAEs, while the third response was defined as integrated MAE.

There was a rest period of at least 45 seconds after each response. Each observer completed 8 trials for each condition in each experiment, with each trial corresponding to one of the 8 integrated directions. Orders of conditions and directions were randomized for each subject.

Stimuli were generated using MATLAB and PsychToolbox (Brainard, 1997; Pelli, 1997) and presented on a Viewsonic CRT monitor with a refresh rate of 75 Hz and resolution of 1024×768 pixels, with a viewing distance of 57 cm kept constant using a chinrest and forehead rest. Undergraduate students at the University of California, Los Angeles (UCLA), all with normal or corrected-to-normal vision and naïve to the experimental purpose, participated in the three experiments for course credit. The experiments were approved by UCLA’s Office for Protection of Research Subjects.

Experiment 1: Comparing perceived MAE direction tested at different locations

Methods

Test locations were manipulated across three conditions. We first tested the MAE at all adapted locations (the *All* condition), as shown in Figure 3.3A (left panel of Test).

We then constructed a critical condition (the *Single* condition), in which the test pattern consisted of elements taken from only one of the two component sets (Figure 3.3A, middle panel of Test). The third condition was the *Mixed* condition (Figure 3.3A, right panel of Test), in which a random half of the test elements from each set were included. This manipulation of the *Mixed* condition served to match the test element density in the *Single* condition, so that the *Mixed* and *Single* conditions only differed in terms of test locations. Ten naïve observers participated in Experiment 1 for course credit.

In the *Single* condition, the two segregated MAE responses were classified as *tested* or *untested* segregated MAEs using the following procedure. Suppose, in a *Single*-condition trial, test elements were chosen from component set 2 with adaptation direction being (X+45)°. We defined the MAE direction (X-135)°, opposite to the tested set's adapting direction, as the *tested* segregated MAE response, and the other segregated MAE direction (X+135)° as the *untested* segregated MAE. In order to provide a fair comparison between the *Single* condition and the *All* and *Mixed* conditions, in the latter two conditions, one of the two component sets was randomly chosen as the "tested set", so that a segregated MAE response could be assigned accordingly to be *tested* or *untested*.

Results

The distinction between *tested* and *untested* segregated MAE responses in the *All* and *Mixed* conditions is for illustration purpose only. In our statistical analyses, we summed over the proportions of the two segregated responses in order to have a more conservative comparison within each of the *All* and *Mixed* conditions. In both the *All* and *Mixed* conditions, the proportion of integrated MAE (*All*: 71.3%; *Mixed*: 77.5%) was much higher than the sum of the two

segregated MAE response proportions (*All*: 16.3%, $F(1, 9) = 27.99$, $p < .001$; *Mixed*: 12.5%, $F(1, 9) = 76.69$, $p < .001$) (Figure 3.3B). Crucially, the pattern was reversed in the *Single* condition: the proportion of tested segregated MAE responses (72.5%) was much higher than that of integrated MAE (23.8%, $F(1, 9) = 19.76$, $p = .002$). These results clearly demonstrate that when test elements are presented at locations of both component sets, an integrated MAE is found, with the aftereffect direction opposite to the average direction. However, when MAE was tested at the locations of only one component set (e.g., Set 1, adapting direction = (X-45)°), a segregated MAE was obtained, with the aftereffect direction opposite to the tested set's global adapting direction (e.g., MAE direction = (X+135)°). Proportion of "no-motion" response was low across all conditions (means for *All* = 12.5%; *Single* = 3.8%; *Mixed* = 10.0%), indicating that relatively strong motion aftereffects were perceived by observers after transparent motion adaptation

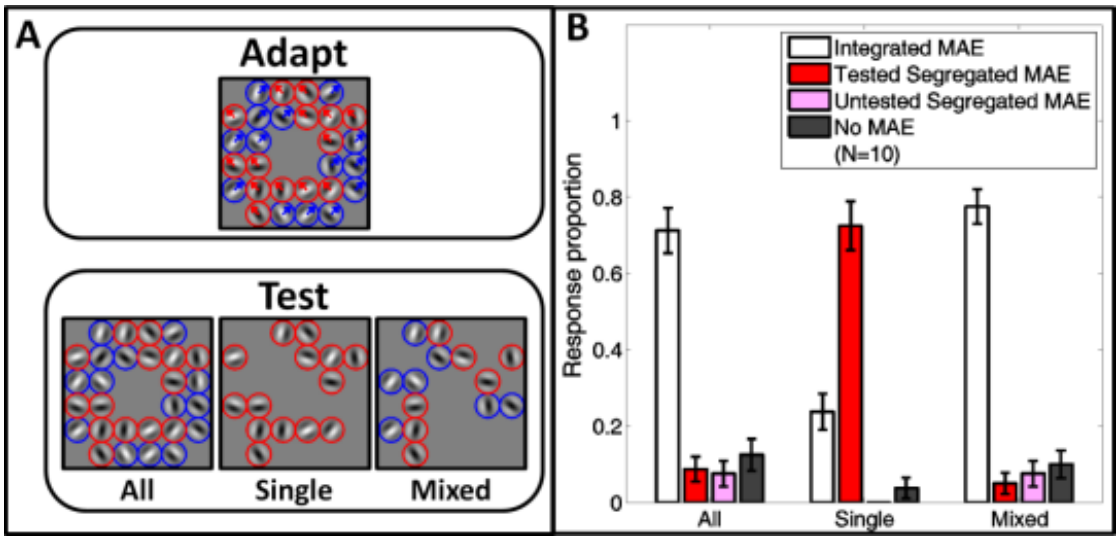

Figure 3.3. Conditions and results of Experiment 1. A: Adapting stimulus was the same across conditions. Test conditions differed only in terms of where test elements were presented. B: Proportions of reported MAE directions. Error bars are 1 S.E.M.

In order to confirm that observers perceived only one MAE direction on each trial, we reran the *Single* and *Mixed* conditions in Experiment 1 using a direction judgment task. Nine fresh observers were instructed to indicate as many directions as they had perceived during the test phase by turning a simulated dial on the screen. As shown in Figure 3.4A, the dominant response for the number of perceived MAE directions was one direction in both the Single and Mixed conditions. In terms of perceived MAE directions (Figure 3.4B), results were similar to that found in the forced-choice task reported above, namely, observers primarily reported the tested segregated MAE direction (opposite to 45° in the figure) in the *Single* condition, and the integrated MAE direction (opposite to 0° in the figure) in the *Mixed* condition.

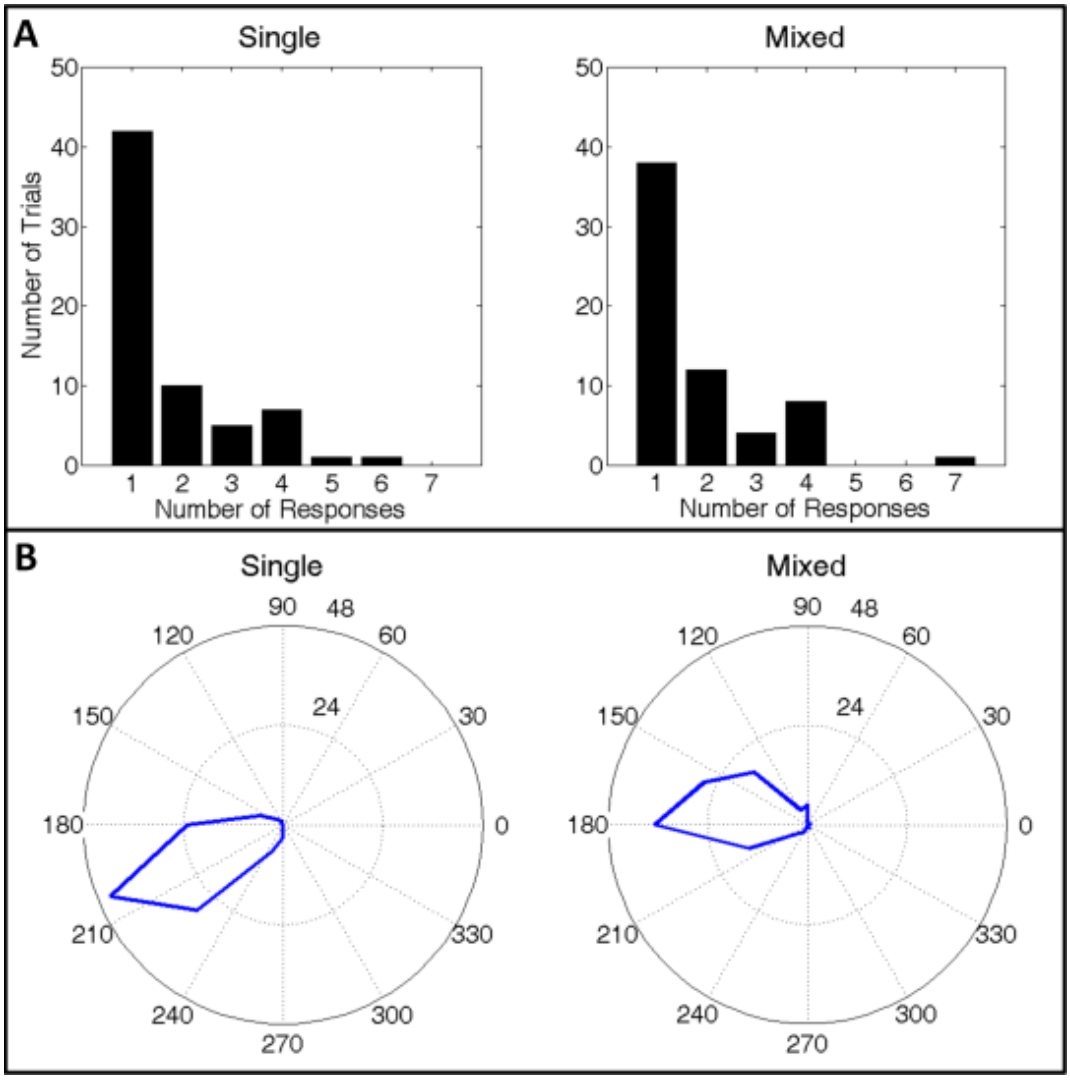

Figure 3.4. Distributions of perceived MAE directions in the dial-turning task for Experiment 1. A: Histograms of the number of responses for the Single (left) and Mixed (right) conditions. B: Distributions of perceived MAE direcitons for the Single (left) and Mixed (right) conditions. Angles represent reported directions. Distance from center represents number of trials. Responses are aligned so that the integrated direction (X°) in the adapting stimulus is at 0°. For the Single condition, responses are aligned so that the tested set's adapting direction is at 45° for every response.

Experiment 1 showed that, by manipulating testing locations, observers can be led to perceive either of two radically different forms of MAE after transparent motion adaptation. However, there are several possible ways in which adaptation could affect different motion processing levels to elicit the two forms of aftereffects. One possible mechanism is that transparent motion adaptation alters response characteristics of global motion processors at MT, which integrates local motion information over the entire display window. As a result, adapting to two global motion directions (X+45)° and (X–45)° in a transparent motion display may change the tuning characteristics of MT cells that prefer the combined direction X°. This mechanism would lead to strong integrated MAE in the direction of (X–45)°, but not location-specific segregated MAE. Another possibility is that the changes introduced at MT are due to the integration of local adaptation effects in early motion processing. This mechanism would predict that perceived MAEs would be determined by the testing locations. However, this possibility hinges on the existence of local adaptation effects: If local aftereffects were eliminated, MAEs would not be observed. A third possibility is that both mechanisms exist in the motion system, as found in previous electrophysiological studies of contrast adaptation (Baccus & Meister, 2004; Kohn & Movshon, 2003; Solomon, et al., 2004), and that they interact to produce motion aftereffects. Experiments 2 and 3 aim to investigate which of the above possible mechanisms underlie the perceived aftereffect of transparent motion.

Experiment 2: Testing MAE at Non-adapted Locations

Methods

Experiment 2 examined whether perceived MAEs depend on integration of local adaptation effects. Previous studies showed that motion aftereffects can be observed at non-adapted regions, a phenomenon termed "phantom" MAE (Snowden & Milne, 1997). The existence of phantom MAE provides evidence that adaptation induces changes at higher-level processing levels with larger receptive fields. In Experiment 2, test locations were fixed, and they were either adapted or non-adapted in the adaptation phase, depending on the condition (Figure 3.5A). Relative to the total number of elements (100%), 75% (37.5% from each of the two sets) were presented during the adaptation phase, leaving 25% (12.5% from each set) as non-adapted. Only 25% of elements were presented during the test phase, regardless of condition. In the *Adapted* condition, adapting elements from the *same* set were presented at test locations. As a result, the *Adapted* condition was identical to the *Single* condition of Experiment 1, except that fewer elements were presented in both adaptation and test phases. In the *Non-adapted* condition, adapting elements were not presented at the test locations during the adapting phase in order to eliminate local adaptation effects. In other words, the test locations were empty during adaptation. Experiment 2 aimed to examine the extent to which the segregated MAEs depend on the propagation of local adaptation effects in the early motion processing level. If the segregated MAE is largely due to integration of local adaptation effects, eliminating local motion adaptation would reduce the likelihood of perceiving segregated MAE. Thirteen naïve observers participated in Experiment 2 for course credits.

Results

Figure 3.5B shows that the tested segregated MAE (60.6%) was more frequently reported than the integrated MAE (22.1%) in the *Adapted* condition ($F(1, 12) = 12.63$, $p < .01$) (Figure 3.5B, left group), replicating the findings of the *Single* condition in Experiment 1. The negligible proportion of untested segregated MAE (1.9%) could be due to response noise (as it was not observed in Experiment 1) and was not considered in the analysis. In contrast, when tested at non-adapted locations (Figure 3.5B, right group), observers still perceived an MAE, but this "phantom" MAE was found to be dominated by the integrated direction (37.5%) as the proportion of segregated MAE responses (15.4%) was smaller ($F(1, 12) = 10.11$, $p < .01$). As we switched from the *Adapted* to the *Non-adapted* conditions by eliminating the effects of location adaptation, the proportion of segregated MAE responses greatly reduced from 60.6% to 15.4% ($F(1, 12) = 41.55$, $p < .001$). This significant reduction in segregated MAE clearly supports the hypothesis that the segregated MAE mainly results from the integration of local adaptation effects. A significant two-way interaction (MAEs × Conditions: $F(1, 12) = 25.86$, $p < .001$) further demonstrates that the elimination of local adaptation shifted the perceived aftereffect direction from segregated MAE to integrated MAE.

Note that the perceived aftereffect in the *Non-adapted* condition was relatively weak. The proportion of "no-MAE" responses was 47.1%, meaning that observers reported perceiving an MAE in only 52.9% of trials (c.f. "no-MAE" proportion in *Adapted*: 15.4%, MAE proportion: 85.6%). The weaker strength of aftereffect at non-adapted locations is consistent with other findings concerning phantom motion aftereffects (Snowden & Milne, 1997). In addition,

observers in our experiment were naïve undergraduate students who participated for course credits. It is likely that the inexperience of subjects also contributed to the weak effect of adaptation. Nonetheless, our results indicate that observers consistently reported integrated MAE when they perceived an aftereffect in a trial in the *Non-adapted* condition. In order to find out whether observers reported integrated MAE more frequently than chance level, we computed the normalized proportion for the integrated MAE responses. For each condition, we divided the integrated MAE response proportion by the total proportion of MAE responses. For a more conservative test, we compared this value with the chance-level performance in a 2AFC task (50%), as if observers were to choose between the integrated and segregated MAE directions. We found that the integrated MAE was perceived more frequently than chance in the non-adapted regions (Normalized proportion of integrated MAE responses = 73.0%; $t(11) = 3.40$, $p = .006$)[4]. This indicates that, to some extent, adaptation at higher processing levels contributes to the perceived phantom MAE in non-adapted locations.

[4] One observer reported "no-motion" in all trials of the *Non-adapted* condition. It is not possible to compute a normalized integrated MAE response proportion for this observer. Although this observer was included in the main analysis, he/she was not included in the test of "normalized proportion vs chance").

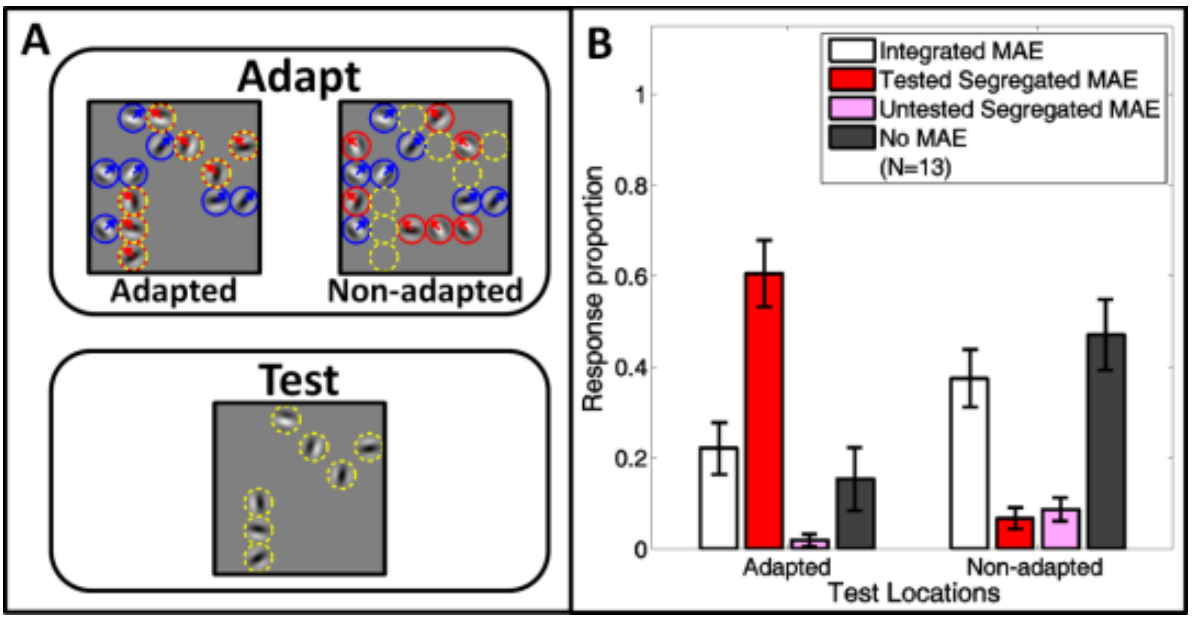

Figure 3.5. Conditions and results of Experiment 2. A: Unlike Experiment 1, test locations were held constant (dashed-yellow circles). The state of adaptation of these fixed test locations was manipulated. In the *Adapted* condition (top-left), test locations overlap with adapting elements taken from one of the two component sets (dashed-yellow on red, in this illustration). In the *Non-adapted* condition (top-right), test locations (empty dashed-yellow circles) did not overlap with any adapting elements. B: Proportions of reported MAE directions. Error bars are 1 S.E.M.

This manipulation in our Experiment 2 is similar to that used in a recent study by Scarfe and Johnston (2011). The key difference is that we used transparent motion as the adapting stimulus, whereas Scarfe and Johnston used unidirectional moving Gabor arrays as the adapting stimulus in their Experiment 3. Participants in their experiment reported "no MAE" on more than 90% of trials, indicating that phantom MAE was not observed in their experiment. While

observers in their experiment were all experienced in psychophysical experiments, ours were all inexperienced observers. It is possible that the ability of inexperienced observers to maintain a steady fixation during adaptation and/or test was poor. Accordingly, eye movements may contribute to the reported phantom MAE, instead of the adaptation of some high-level mechanisms.

We assessed this possibility with a follow-up experiment using similar procedure as the *Non-adapted* condition in Experiment 2, but employed a different adaptation stimulus and a different task. Instead of two translational 2D vectors assigned to the two component sets of elements, global motion for each set was a spiral (radial + rotational). Elements in one set were assigned 2D velocities that were consistent with a global clockwise-inward (or outward, radial directions randomized across trials) spiral, which is analogous to the (X+45)° direction in the original experiments. Elements in the other set were assigned a counterclockwise-inward (or outward) spiral, which is analogous to the (X-45)° direction. This stimulus was perceived to be an overall radial motion, with a certain level of rotational transparency. After adapting to this stimulus, observers were presented with the static test stimulus, with all elements shown at non-adapted locations (same as the *Non-adapted* condition in Experiment 2). During the response phase, observers first indicated whether they saw an MAE or not. If they indicated that they saw an MAE, they were then asked to indicate the direction of the perceived MAE in a 3AFC task, with the 3 alternatives being clockwise-outward spiral, purely outward and counterclockwise-outward spiral (for inward adapting stimuli; responses were "inward" if adapting stimuli were outward). If they reported "no MAE", the trial ended at that point. Four experienced observers (at least 3 years of experience in visual psychophysics), who

knew neither the purpose nor the design of the experiment, participated. Each observer completed 16 trials.

Results from individual observers are shown in Figure 3.6. All 4 observers consistently reported seeing the phantom MAE above chance level, except that JS reported the phantom MAE in 10 out of 16 trials, slightly above chance (Figure 3.6A). To assess how frequently each MAE direction observers perceived, we computed the normalized response proportion for each MAE direction. The proportions were "normalized" against the total number of trials in which MAE was perceived (e.g., if JS perceived a purely radial MAE in 8 trials out of the 10 trials in which she saw an MAE, the normalized proportion for "radial" would be 80%). There is a consistent trend across observers (Figure 3.6B): the dominant perceived direction of the phantom MAE was in the radial direction, which is analogous to the dominant "integrated" direction in the phantom MAE described in Experiment 2.

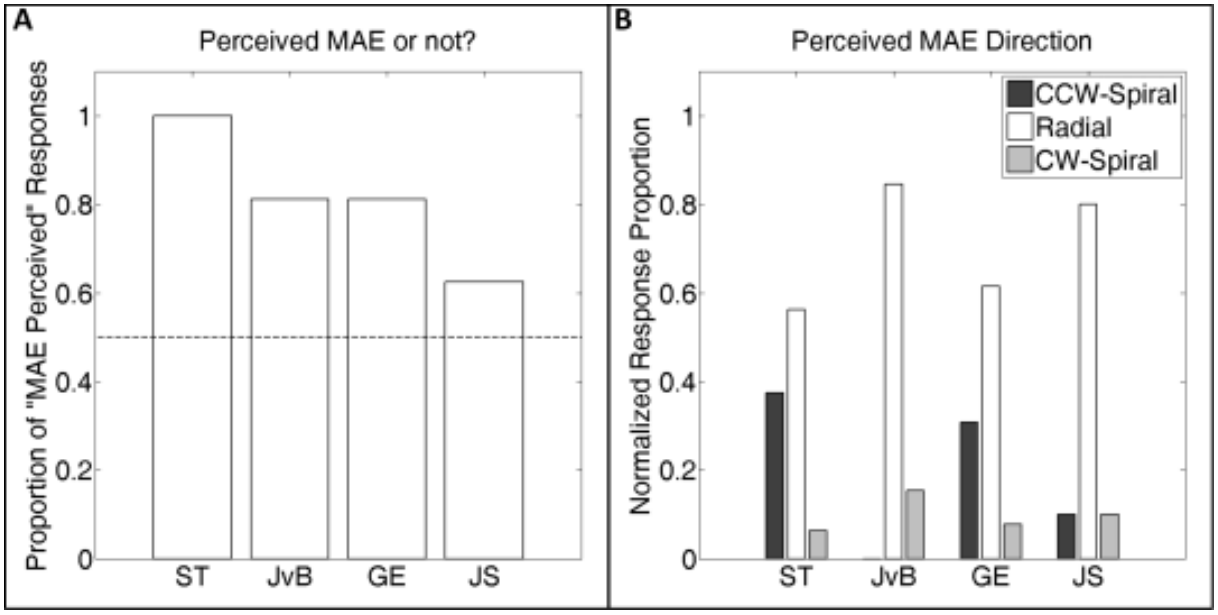

Figure 3.6. Results of individual experienced observers on complex phantom MAE. A: Proportion of trials in which observers indicated that they saw MAE. B: Response proportions for each directional response, normalized within trials in which MAE was perceived.

Results from the follow-up experiment are consistent with those in Experiment 2, providing further evidence for the existence of an integrated, phantom MAE after adapting to transparent motion generated from a multiple-aperture stimulus. The results also show that it is unlikely that the phantom MAE can be explained purely by eye movements. This finding is inconsistent with those reported by Scarfe and Johnston (2011). We suspect that this discrepancy is mainly due to the difference in adaptation durations. In their study, adaptation lasted for 8 seconds on the first trial, with 3-second top-ups. An 8-second adaptation is much shorter than the durations used in previous phantom MAE studies, in which adaptation duration ranged from 30 to 120 seconds, with

top-ups of 4 to 5 seconds (Price, Greenwood, & Ibbotson, 2004; Snowden & Milne, 1997; Weisstein, Maguire, & Berbaum, 1977) . In our experiments, adaptation lasted for 45 seconds on seconds on all trials, which is within the range of adaptation duration used in previous studies in which phantom MAE has been observed. It has also been shown that the strength of MAE positively correlated with adaptation duration (Hershenson, 1993). Another potential reason for the discrepancy may involve contrasts. Our stimulus employed low contrast (0.05) for both adapting and testing stimuli, whereas Scarfe and Johnson (2011) used 0.3 contrast for adapting and 0.8 for testing stimuli. As suggested by Amano et al. (2009), high stimulus contrasts may impair spatial integration, which could weaken the percept of global motion generated from a multiple-Gabor stimulus. We conjecture that low contrasts may also strengthen motion integration and interpolation after adaptation, which in turn increases the possibility of perceiving the illusory phantom MAE.

Experiment 3: Testing MAE at Non-adapted orientations

Methods

Experiment 3 varied local adaptation effects by exploiting the property that local-motion processors in V1 are orientation-selective. Previous studies (Hammond, Pomfrett, & Ahmed, 1989; McGraw, Whitaker, Skillen, & Chung, 2002) have shown that motion aftereffects are substantially reduced when the test grating has an orientation orthogonal to that of the adapting grating. We therefore manipulated the orientations of test elements, either keeping them the same as adapting elements or rotating them by 90° from adapting orientations. As in Experiment 1, we included testing locations as an independent variable. Accordingly, this design involved four

experimental conditions (Figure 3.7A): *Mixed*, *Orthogonal-Mixed*, *Single*, and *Orthogonal-Single*, which allows us to examine how local adaptation effects interact with integrated MAE. Nineteen naïve observers participated in Experiment 3 for course credit.

Results

As shown in Figure 3.7B, when test elements are presented at locations of both adapting sets (the two *Mixed* conditions), integrated MAE was more frequently reported than segregated MAEs in both conditions (*Mixed*: Integrated (59.9%) vs. sum of segregated (19.1%), $F(1, 18) = 34.46$, $p < .001$; *Ortho-Mixed*: Integrated (43.4%) vs. sum of segregated (19.7%), $F(1, 18) = 11.83$, $p < .005$). However, this dominant integrated MAE was weakened when test orientations were made orthogonal to adaptation orientations. The proportion of integrated MAE in the *Orthogonal-Mixed* condition (43.4%) was lower than that in the *Mixed* condition (59.9%), $F(1, 18) = 5.580$, $p < .05$. This finding indicates that integration of local adaptation effects may contribute to the integrated MAE. When test locations were chosen from only one adapting set, a dominant tested segregated MAE (58.6%) was obtained when identical orientations were used for adapting and testing stimuli (*Single:* tested segregated (58.6%) vs. integrated MAE (27.6%), $F(1, 18) = 9.48$, $p < .01$), replicating the results in Experiment 1. As in Experiment 2, the negligible proportion of untested segregated MAE responses, 2.0%, was not included in the analysis. However, when orthogonal orientations were tested in the *Orthogonal-Single* condition, perceived MAE was biased more toward the integrated (45.4%) than the segregated directions (tested = 15.1%; untested = 7.2%). Specifically, the proportion of integrated MAE responses was larger than the sum of the two segregated MAE response proportions ($F(1, 18) = 9.32$, $p < .01$). A significant simple

two-way interaction between test orientation and perceived MAE direction for the two *Single* conditions ($F(1, 18) = 18.62$, $p < .001$) provides further evidence that the integration of local adaptation effects was the major mechanism contributing to the perceived segregated MAE. In addition, there were generally more "no-motion" responses in the orthogonal conditions (*Orthogonal-Mixed* = 36.8%; *Orthogonal-Single* = 32.2%) than in the same-orientation conditions (*Mixed* = 21.1%; *Single* = 11.8%), $F(1, 18) = 11.03$, $p < .005$, confirming that the use of orthogonal test orientations effectively weakened the strength of perceived motion aftereffects.

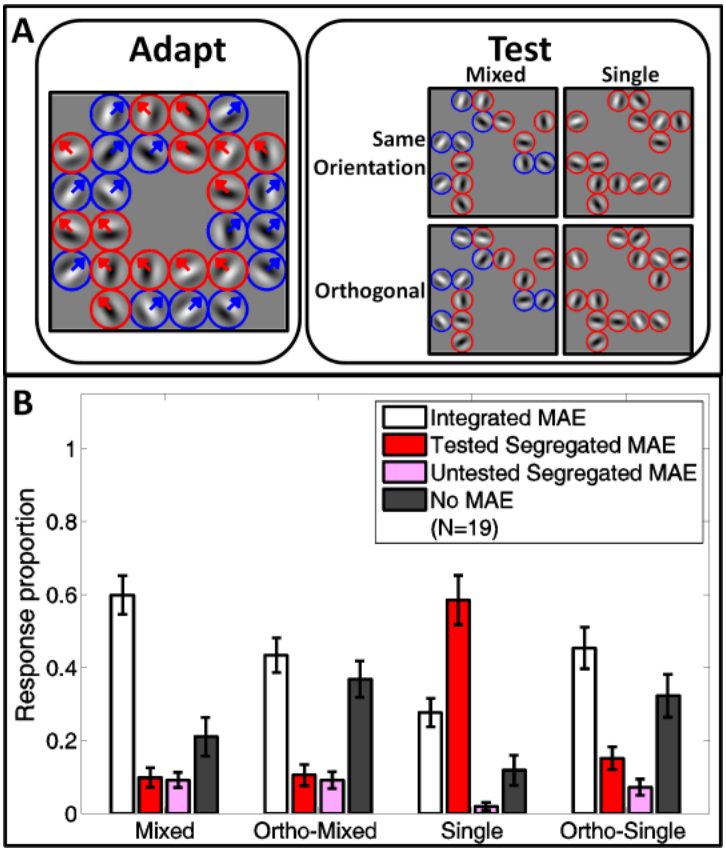

Figure 3.7. Conditions and results of Experiment 3. A: Procedure and conditions were the same as in Experiment 1, except for the introduction of test orientation as an independent variable, resulting in 4 conditions. B: Proportions of reported MAE directions. Error bars are 1 S.E.M.

General Discussion

The present study shows that adaptation to bidirectional transparent motion can induce two radically different motion aftereffects, and elucidates the linkage between these perceptual aftereffects and different motion adaptation mechanisms in the visual system. Experiment 1 shows that test locations determine which form of MAE observers perceive after adapting to transparent motion. When the test stimulus contained elements chosen from only one of the two component sets, perceived MAE direction was found to be opposite to adapting direction of the chosen set (a *tested segregated MAE*). When the test stimulus contained elements randomly chosen from both sets, perceived MAE direction was opposite to the average of the two adapting directions (an *integrated* MAE). Experiment 2 investigated the mechanisms underlying the two forms of MAE. We found that, when local adaptation was eliminated, segregated MAE diminished and integrated MAE reemerged. This finding suggests that the segregated MAE results from integration of local motion aftereffects via the propagation of adaptation-induced changes in early processing areas (i.e., V1), and part of the integrated MAE is due to global motion adaptation via directly modulating responses of neurons at the global processing stage (i.e., MT). Experiment 3 shows that integration of local motion aftereffects, as a second source of adaptation-induced changes, also plays a role in determining the magnitude of integrated MAE perceived by the observers.

The multiple-aperture adapting stimulus used in our experiments is able to yield the percept of bidirectional motion transparency. Amano et al. (2009) reported that, when Gabor elements in the same component set were assigned the same orientation, the "transparent" pattern was perceived to be unidirectional. However, by randomizing the

orientations of elements within each component set, we increased the variability in local drifting velocities and, thus, reduced the coherence across the two sets. This made the two individual directions of the components more salient. As shown in the results reported of our preliminary experiment, observers simultaneously perceived the two global directions when presented with our adapting stimulus, instead of perceiving the unidirectional, integrated one.

Our general findings are consistent with those reported by Scarfe and Johnston (2011). Both studies indicate that the visual system enables the integration of ambiguous local motion signals to infer global motion, and, meanwhile, retains local spatial precision that is important for representing motion boundaries and features (O. Braddick, 1993).These findings are consistent with a line of research on the importance of local adaptation in generating a MAE percept (Curran, et al., 2006; Lopez-Moliner, Smeets, & Brenner, 2004; Vidnyanszky, et al., 2002). However, a significant discrepancy between the two studies concerns findings of phantom MAEs using multiple-aperture stimulus. We found evidence of a weak phantom MAE at non-adapted locations after adapting to transparent motion, whereas Scarfe and Johnson's (2011) study did not reveal any phantom MAE after adapting to unidirectional motion. We conjecture that this difference may be largely due to adaptation duration and stimulus parameters, such as contrasts.

Our results show that transparent motion adaptation induces changes at both local and global levels of motion processing. Two different neural adaptation mechanisms are involved: integration of adaptation effects at the local processing level and adaptation-induced modulation at the global level. The hypothesis that the visual system utilizes a combination of these two mechanisms can explain why most previous studies on transparent motion adaptation found an integrated MAE but not a segregated MAE, which was found in our *Single* conditions. A multiple-aperture stimulus enhances the possibility of finding the two forms of aftereffects,

mainly due to orientation selectivity of local motion detectors in V1. Random-dot (Alais, Verstraten, & Burr, 2005; van der Smagt, Verstraten, & van de Grind, 1999) or random-pixel (F. A. Verstraten, et al., 1994) patterns contain a distribution of orientations at each location, and stimulate detectors preferring a range of orientations within one location. Therefore, it is challenging to design a test stimulus for these adapting patterns to examine the specific local aftereffects generated by orientation-selective local motion detectors. In contrast, the multiple-aperture stimulus specifically stimulates local motion detectors at one orientation at each location. This stimulus makes it possible to design a test stimulus such that effects of adaptation at those adapted orientations (or non-adapted orientations, as in Experiment 3) can be measured. Such differences in the potential to extract local MAEs between types of stimuli explain why most previous studies did not find the segregated MAE.

In the psychophysics literature, the segregated MAE has never been observed after transparent motion adaptation, unless other cues distinguish the two component motion patterns, such as speed (van der Smagt, et al., 1999; F. A. Verstraten, van der Smagt, Fredericksen, & van de Grind, 1999), temporal frequency (Alais, et al., 2005) or binocular disparity (Curran, Hibbard, & Johnston, 2007; F. A. J. Verstraten, Verlinde, Fredericksen, & Vandegrind, 1994). However, the segregated MAEs observed in these previous studies could possibly result from two different neural adaptation mechanisms. (1) Adaptation may induce changes at a low-level processing stage, and then local aftereffects associated with the same cue value may be integrated to produce the perceived segregated MAE. (2) Adaptation may induce changes in different tuning channels (e.g., speed-tuned channels in MT) at the global motion processing level.

Accordingly, the paradigms used in these previous studies could not clarify which adaptation mechanism the visual system adopts. In our study, spatial separation allows us to probe the adaptation to each component motion after adapting to a transparent motion stimulus. Although spatial separation of component motions accomplishes practically the same function as assigning different speeds to two component motions, there are no tuning channels at the global motion processing level corresponding to different spatial separations. The segregated MAEs revealed in our study thus provide direct evidence of the integration of local adaptation effects and the mechanisms of multilevel adaptation.

Our results show that adaptation is a complicated process, and an observed aftereffect is not necessarily generated from one single neural site. With an adapting stimulus that stimulates multiple levels of processing, observed aftereffects could result from neural changes at one level, propagated effects from lower to higher levels, and/or a combination of the two. Future research on sensory adaptation should focus on how to tease apart the contributions of multiple possible mechanisms for adaptation-induced changes in neural processing and perception.

Chapter 4

A Motion Aftereffect purely generated by Local Adaptation

Introduction

Prolonged visual stimulation leads to biased perception on the subsequent stimulus. These robust perceptual effects of adaptation, collectively known as aftereffects, reveal the perceptual consequence of neural adaptation in the visual system in response to recent stimulus history (C. W. Clifford, et al., 2007; Kohn, 2007; Webster, 2011). Adaptation, therefore, has been widely used as "psychophysicists' microelectrode" (Thompson & Burr, 2009) to examine specific mechanisms in the visual system. However, the process by which neural adaptation leads to perceptual aftereffects remains unclear. In particular, as visual analysis involves multiple stages of processing, how adaptation at early processing stages propagates to produce perceptual aftereffects is still under investigation.

Vidnyanszky and colleagues (2002) proposed that the perceived direction of a global motion aftereffect (MAE) can be computed by spatially integrating local, illusory motion signals. This proposal is consistent with the hypothesis of propagation of adaptation effects from low-level to high-level motion processing, which is supported by findings from single-cell recording (Kohn & Movshon, 2003, 2004), neuroimaging (H. A. Lee & Lee, 2012), as well as psychophysical experiments (Curran, et al., 2006; Lee & Lu, 2012). Similar patterns of results have been observed in a study of face/curvature adaptation (Xu, et al., 2008), suggesting that such a bottom-up propagation mechanism

may be a general adaptation strategy employed by the visual system . Nonetheless, what computation underlies this mechanism still remains an unanswered question.

The challenge in examining this propagation mechanism, namely, *integration of local illusory signals*, is amplified by the difficulty of dissociating it from another adaptation mechanism, *adaptation at the global motion processing stage*. Prolonged visual adaptation generates illusory signals at local locations, which are automatically integrated over space via some global, integrative mechanisms. However, during the course of adaptation, these global mechanisms are also activated, adapt, and, thus, contribute to the generation of MAE. In order to examine the two mechanisms separately, some attempts have been made to study global motion adaptation by weakening adaptation at the local-processing stage, e.g., employing limited lifetimes in the random-dot kinematograms (RDKs) display (e.g., Grunewald & Lankheet, 1996; F. A. J. Verstraten, et al., 1994), and testing MAEs at non-adapted regions or locations (e.g., Scarfe & Johnston, 2011; Snowden & Milne, 1997). However, little work has been done on suppressing global adaptation to focus on the contribution of the propagation mechanism via integration of local illusory signals.

One possible way to suppress global adaptation is to mask the stimulus from conscious perception while maintaining low-level access to certain stimulus properties during adaptation. Using various techniques, such as spatial crowding (He, Cavanagh, & Intriligator, 1996; Rajimehr, Vaziri-Pashkam, Afraz, & Esteky, 2004; Whitney & Bressler, 2007), binocular suppression (Blake, Tadin, Sobel, Raissian, & Chong, 2006; Maruya, Watanabe, & Watanabe, 2008) and tuning the stimulus parameters to imperceptible levels (He & MacLeod, 2001), psychophysicists have successfully prevented observers from consciously perceiving certain aspects (e.g., orientation, motion direction) of the adapting stimulus. Interestingly, significant

aftereffects were found in these studies even when observers were unaware of the characteristics (sometimes even the presence) of the stimulus during adaptation.

The present study applied a similar method to suppress adaptation at the global motion processing stage. We used a particular adapting stimulus, which would not evoke a conscious percept in terms of coherent motion direction. For motion perception, it has been shown in physiological studies with primates (Bradley, Chang, & Andersen, 1998) and neuroimaging studies (Tootell, et al., 1995) that conscious perception appears to correlate with activity in high-level global-motion-sensitive neural units (e.g., MT in non-human primates or hMT+ in humans). To suppress MT activity at the global motion processing stage, motion directions can be made imperceptible using multidirectional stimuli. Treue and colleagues (2000) demonstrated that, by manipulating the number of global motion directions, it is possible to create two motion stimuli that are physically different but perceptually indistinguishable. This method was later extended in other psychophysical studies (Edwards & Greenwood, 2005; Greenwood & Edwards, 2009), which demonstrated that humans are unable to discriminate motion patterns composed of 4 global directions from patterns including 5 directions. The present study employed multidirectional motion stimuli to mask conscious perception of global motion directions during adaptation. This would allow us to isolate the *integration of local illusory signals* from the *adaptation of the global-motion processing stage*, and focus on the contribution of the former in producing MAE. If observers can perceive an aftereffect after adapting to this imperceptible, multidirectional pattern, it would suggest that the visual system maintains and utilizes local motion information without conscious awareness, despite its inherent ambiguity due to the aperture problem. Furthermore, these motion processors

can adapt and generate local illusory motion signals, which will be propagated via spatial integration to generate the perceived MAEs.

We conducted a psychophysical and computational investigation of how adaptation induces changes at the local motion stage, and how these changes propagate to generate perceptual aftereffects. First, we report the results of two psychophysical experiments. We measured the perceived direction of an MAE that was generated purely based on integration of local illusory signals. In Experiment 1, the adapting stimulus consisted of 5 global translational directions within the same spatial region; in Experiment 2, the adapting stimulus consisted of 4 complex (circular and radial) global directions. Second, we present a physiologically-inspired model that produces and integrates adaptation-induced local illusory signals. The performance of the model demonstrates that such integration of local illusory signals *per se* is sufficient to produce global motion aftereffects.

General Methods

Stimulus

The stimulus consisted of 264 drifting Gabor elements, as shown in Figure 4.1A. Each element was a sine-wave grating, windowed by a stationary Gaussian function with a sigma of 0.21°. Spatial frequency was kept constant at 2 cycles/°. Each element subtended a visual angle of 1°. Orientations were randomly and independently sampled from a uniform distribution between 0° and 180°. Elements were arranged in a circular pattern inscribed in a 20×20 grid in which cells were tightly packed, so that separation between any two adjacent cells was zero. As a result, the centers of any two adjacent Gabors were 1° apart. Elements that were too close or too far away from fixation were removed, resulting in a display annulus spanning 4°–10° around fixation. Contrast was fixed at a low level of 0.05 for all experiments, in order to facilitate the perception of global motion based on spatial integration of local motion signals in the multiple-Gabor stimulus.

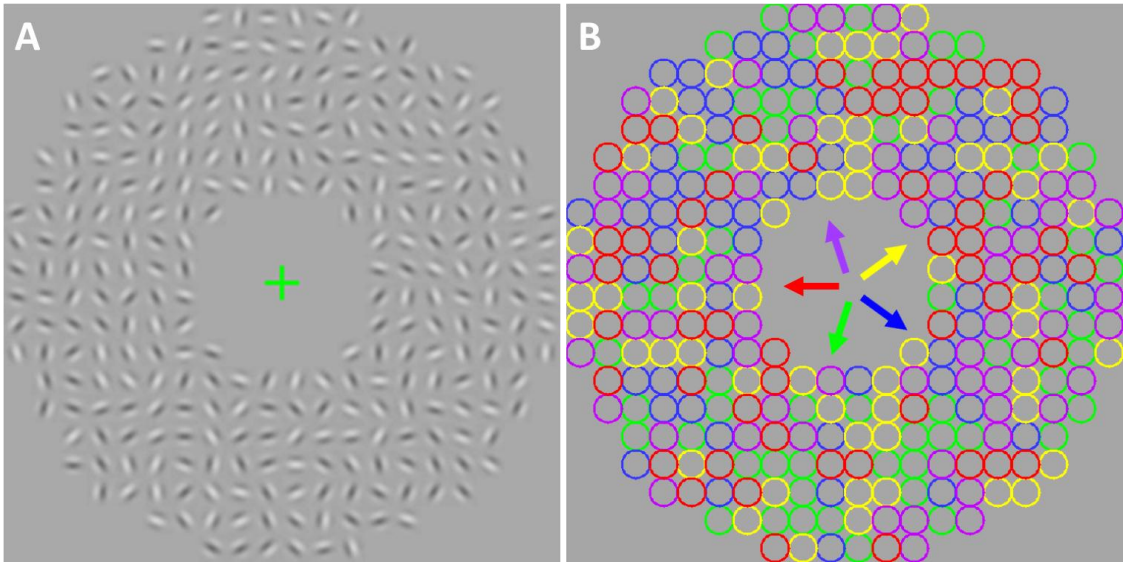

Figure 4.1. Illustration a 5-direction stimulus. Parameters (e.g., contrast) have been modified for demonstration purpose. (A) Layout of the multiple-Gabor pattern. (B) Schematic illustration of 5-direction motion pattern. Colors indicate different sets of Gabor elements. Circles indicate locations of elements. Arrows at the center indicate global velocities assigned to different sets.

Elements were randomly assigned to different sets (2, 4 or 5 sets, depending on experimental conditions; Figure 4.1B illustrates a 5-set pattern). Each set was assigned a global motion velocity, with speed = v and direction = α. Local drifting speed u of each element was computed based on the element's orientation θ and the global velocity assigned to the set to which the element belonged:

$$u = v \sin(\alpha - \theta)$$

This assignment was designed to ensure that local drifting velocities of elements within the same set were all consistent with the set's global velocity. In other words, the set's global velocity was the intersection-of-constraint (IOC) solution when combining local motion signals from all elements within this set.

General procedure

We used a "top-up" paradigm to measure motion aftereffects. Specific parameters differed across experiments, but the general procedure remained the same. There were three phases in each trial: adaptation, test, and response. In the adaptation phase, observers viewed a long adapting motion pattern. Observers were instructed to maintain fixation at the fixation cross, which was located at the center of the screen. Before the adapting stimulus ended, three beeps, one second apart, were presented to remind observers about the upcoming test phase. The test stimulus was presented right after the adapting stimulus, and there was no blank frame between the adaptation and the test phases. To avoid abrupt changes from adaptation to test, the last frame of the adapting movie was used as the test image. In the response phase, observers were instructed to make a response after the test stimulus had disappeared, and were encouraged to respond within 3 seconds. Response from the observer triggered the start of the adaptation phase of the following trial. Trials were presented in blocks, with the first trial having longer adaptation duration and the remaining top-up trials shorter adaptation duration. The adapting stimulus of each top-up trial was an excerpt (first frame randomly determined) of the one used in the initial trial within the same block, so that orientations and local motion velocities were identical across all trials in the same block. There was a rest period between every two blocks.

Stimuli were generated using MATLAB and PsychToolbox (Brainard, 1997; Pelli, 1997), and presented in a dim room on a Viewsonic CRT monitor, with refresh rate of 75 Hz and resolution of 1024×768 pixels. Viewing distance was kept constant at 57 cm using a chinrest and forehead rest, so that each pixel on the screen subtended 2.01 arcmin.

We used a Minolta CS-100 photometer to calibrate the monitor, and converted a luminance range of 0–146.5 cd/m^2 into a linear lookup table for 256 intensity levels. Participants were undergraduate students at the University of California, Los Angeles (UCLA), participating for course credit. All participants had normal or corrected-to-normal vision and were naïve to the purpose of the experiments. The experiments were approved by UCLA's Office for Protection of Research Subjects.

Experiment 1: Adapting to imperceptible motion directions yields perceptible MAE directions

The objective of Experiment 1 was to examine whether the visual system can produce a global, directional MAE purely based on integration of local aftereffects. In the experiment, we measured the perceived global MAE direction after adapting to a motion stimulus with five global directions embedded within the same spatial region. In the test phase, we specifically selected a set of locations (the tested set) with adapting velocities that had been consistent with only one of the five adapting directions (similar to the technique used in Lee & Lu, 2012; Scarfe & Johnston, 2011). If observers perceived a global MAE direction opposite to tested set's adapting direction, it would imply that integration of local aftereffects *per se* was sufficient to produce a global directional MAE.

Stimulus

Before conducting the MAE experiment, we wished to confirm that humans cannot identify the global motion directions embedded in a 5-direction motion pattern presented in a multiple-aperture stimulus. In a pilot experiment (n = 11), each observer ran 40 trials, each of which contained a randomly generated multidirectional pattern. In half of the trials, elements in

the pattern were randomly grouped into two sets (the Two-set condition). In the other half of the trials, elements were randomly grouped into five sets (the Five-set condition). Order of conditions was randomly interleaved across trials. In both conditions, all global directions were evenly spaced over the 360° range. In the Two-set condition, the two directions were separated by 180°, and, in the Five-set condition, the five directions separated by 72°. For both conditions, we presented a unique set of directions in every trial by adding a constant perturbation to all the directions. The purpose of showing a unique set of directions for each trial was to remove the effect resulting from cardinal-direction preference. After viewing the stimulus for 30 seconds, observers were instructed to report as many global directions as they had perceived, by turning a simulated dial on the computer screen using the computer mouse.

Results from this pilot experiment (Figure 4.2) show that, while observers were able to clearly identify both directions in the Two-set condition, they were unable to reliably identify any of the five directions in the Five-set condition. In addition, from the distributions of the number of reported directions, it is clear that observers saw two directions most of the time in the Two-set stimulus, but they did not perceive that there were five directions in the Five-set stimulus. This result demonstrates that the global motion directions embedded in the Five-set motion pattern are, indeed, imperceptible.

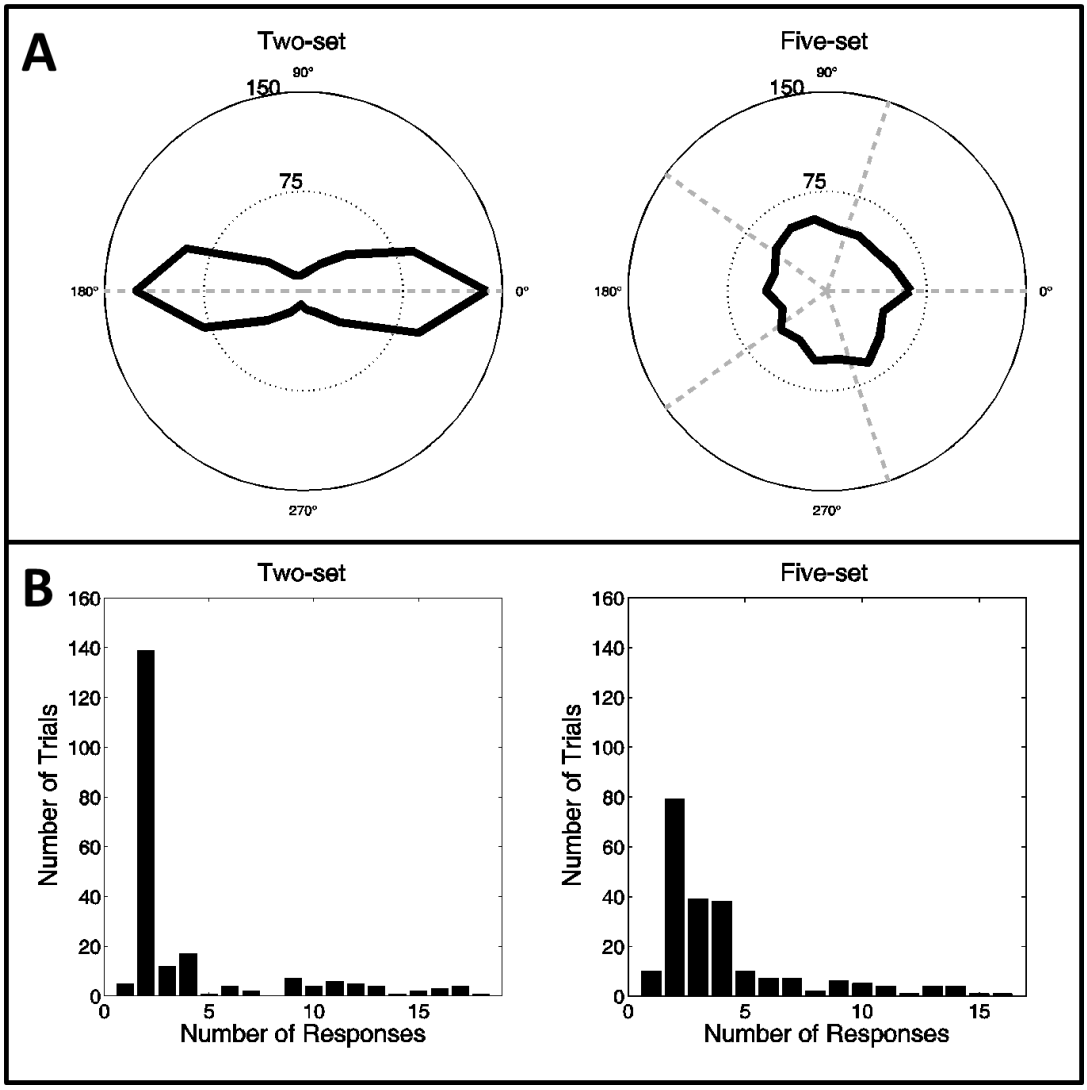

Figure 4.2. Distributions (based on responses from all observers) of perceived motion directions on multidirectional motion patterns with different number of embedded global directions. A) Polar distributions of perceived directions for the two conditions. Angles (small numbers, in degrees) and radial distances (large numbers) represent reported MAE directions and number of trials, respectively. Responses are aligned so that the perturbation for every trial is at 0°. Light dashed lines indicate the embedded directions, and dark solid lines indicate frequencies

for different reported directions. B) Distributions of number of reported directions for the two conditions.

We then used this imperceptible, five-direction motion pattern as the adapting stimulus, and measured the perceived MAE direction under two conditions. Similar to our previous work (Lee & Lu, 2012), the two conditions, namely, the Single and the Mixed conditions, differed in terms of test locations. In the Single condition, all test elements were chosen from the same set. In other words, they all had been assigned the same global adapting velocity in the adaptation phase. This method of testing the MAE has been shown to produce a "segregated" MAE direction, which is opposite to the tested set's adapting direction. In the Mixed condition, one-fifth of the elements from each set were chosen to be presented during the test, so that the total number of test elements was equal in the two conditions. The two conditions were manipulated between subjects. Twenty-five observers (Single: n = 13, Mixed: n = 12) participated in this experiment.

Procedure

Initial and top-up adaptation durations were 30s and 15s, respectively. Test stimulus was a static pattern with elements chosen from different locations depending on the condition, and was presented for 2s. Unlike in the pilot experiment, the task was to report *one and only one* MAE direction perceived during the test phase. Responses were made by turning a simulated dial on the computer screen using the mouse, as in the pilot experiment. Each participant completed ten blocks with a 30-second rest period in between blocks. Each block contained five trials (1 initial + 4 top-ups).

Similar as the Five-set condition in the pilot experiment, ten angles of perturbation, evenly spaced within the 72° range, were randomly assigned to the ten blocks. In the Single condition, one of the five sets of elements was chosen to be the tested set in each trial. Within a block, each set was tested exactly once, so that the five trials in each block would cover all five sets. Test order of the sets was randomized within each block. In each trial of the Mixed condition, one-fifth of elements were randomly chosen from each set to be the test elements. For both conditions, test locations varied across trials within the same block, so that the same location would be tested exactly once in each block. Observers completed two blocks of five practice trials (initial adaptation, top-up adaptation and rest period were 20, 10 and 20 seconds, respectively) before running the real experiment session, in order to familiarize themselves with the procedure.

Results

As shown in the left panel of Figure 4.3, when test elements were taken from the same set in the Single condition (left panel of Figure 4.3), distribution of reported MAE directions was found to be significant deviated from a circular uniform distribution (Hodges-Ajne test, $p < .0001$). The distribution shows a clear peak at 180°, which is opposite to the tested set's adapting direction. However, when test elements were taken from all five sets in the Mixed condition (right panel of Figure 4.3), distribution of reported MAE directions was not different from a circular uniform distribution (Hodges-Ajne test, $p = .687$).

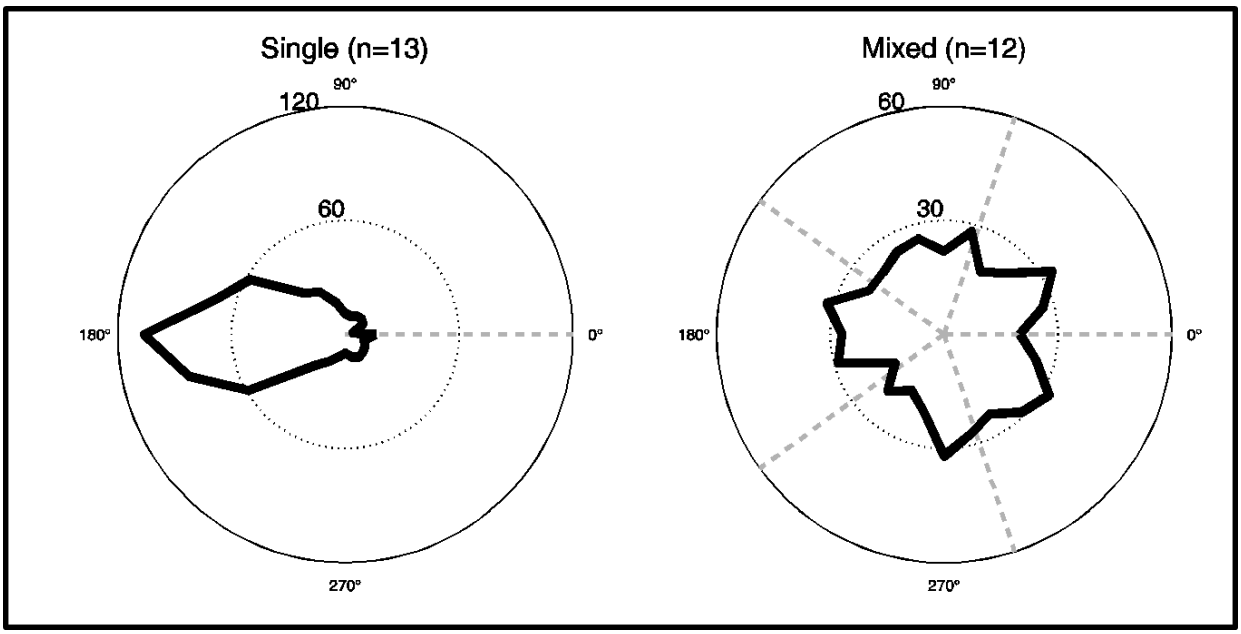

Figure 4.3. Distributions of perceived MAE directions for the Single (left) and Mixed (right) conditions in Experiment 1. Responses are aligned based on the following criteria: for the Single condition (left), 0° represents the tested set's adapting direction for every trial; for the Mixed condition (right), 0° represents the angle of perturbation for every trial. Light dashed lines indicate the adapting directions, and dark solid lines indicate frequencies for different reported directions.

Results from Experiment 1 suggest that, if MAE is tested at locations at which local adapting motion signals are consistent with a coherent global direction, adaptation to imperceptible global directions can lead to perceptible MAE in opposite directions. However, if test locations contain adapting motion signals that are not coherent, there is no clear MAE directional percept. This result implies that the visual system integrates local illusory signals across test locations to general perceived motion aftereffect. Such integration is sufficient to

produce a clear directional MAE, even when the observer is unaware of any coherent motion directions during adaptation.

So far, we have demonstrated the existence of a mechanism that allows the visual system to 1) integrate local illusory signals, and 2) generate perceptible MAEs after adapting to multiple, imperceptible motion directions. However, the motion directions embedded in the adapting stimulus were all unidirectional, which may not necessarily involve levels of motion processing that are beyond the MT level. How far can the integration of local illusory signals reach in the motion processing hierarchy? Experiment 2 was designed to address this issue, by using complex motion patterns (e.g., radial and rotational motion) as the adapting stimulus, in order to involve higher level of motion processing.

Experiment 2: Adapting to imperceptible complex motion patterns yields MAE

Previous findings from neurophysiology (Duffy & Wurtz, 1991a; Tanaka & Saito, 1989), neuroimaging (Morrone, et al., 2000), psychophysics (Lee & Lu, 2010; Morrone, et al., 1995; Snowden & Milne, 1997), and computational models (Beardsley, Ward, & Vaina, 2003; Wu, et al., 2009) suggest that there exist high-level computational processes that are specialized for analyzing complex motion patterns (e.g., optic flow and rotational patterns). They are complex because, unlike unidirectional flows to which the MT area is sensitive, their flow fields contain different local velocities across locations. It is believed that these complex patterns are processed by neurons located in areas that are beyond MT and further downstream in the motion pathway, including the medial superior temporal (MST) area (Duffy & Wurtz, 1991b; Tanaka, et al., 1989). In order to investigate whether the propagation mechanism identified in Experiment 1 extends to the level of complex motion analysis, Experiment 2 employed an adapting stimulus that was

embedded with multiple complex motion directions (rotational and radial motion directions). If the MAE found in the Single condition of Experiment 1 could be observed in Experiment 2, it would suggest that local illusory signals are also integrated at this high-level stage of complex motion processing. However, if integration of local illusory signals is not performed at such a high level of processing, there should be no difference in perceived MAE directions between the Single and Mixed conditions in Experiment 2.

Stimulus

The adapting stimulus was generated in a similar way as in Experiment 1, except that elements were randomly grouped into four sets. Each set was assigned one of the four complex global motion directions – clockwise, counterclockwise, expansion and contraction. Observers were not aware of the embedded motion directions in the adapting stimulus. There were five test conditions. Four of the five test conditions were essentially similar to the Single condition in Experiment 1, in each of which all test elements were taken from the same set. In other words, in a particular trial of any of these four conditions, all test elements were assigned the same complex motion direction in the adaptation phase. We labeled each of these four Single conditions using the tested set's adapting direction: Clockwise, Counterclockwise, Expansion, and Contraction. The remaining condition was the Mixed condition, in which one-fourth of test elements from each set were presented simultaneously during the test phase. In short, the five conditions differed in terms of the complex global motion direction exhibited by the test elements' motion during the adaptation phase.

Procedure

The procedure was similar to that in Experiment 1, except that all five conditions (four Single and one Mixed conditions) were run within-subjects. Each observer ran ten trials for each condition. Order of conditions was randomized across all 50 trials, which were blocked into ten blocks of five trials. Adaptation and test durations were identical to those used in Experiment 1. The task in Experiment 2, unlike that in Experiment 1, was a forced-choice task, in which observers were instructed to report the perceived MAE direction during the test phase by choosing one of the four complex motion directions: CCW (counterclockwise), CW (counterclockwise), Con (contraction) and Exp (expansion). Observers were provided a fifth option to choose "no-motion" (labeled as "NoMAE" in Results) if no MAE was perceived during the test phase. Ten observers participated in this experiment.

Results

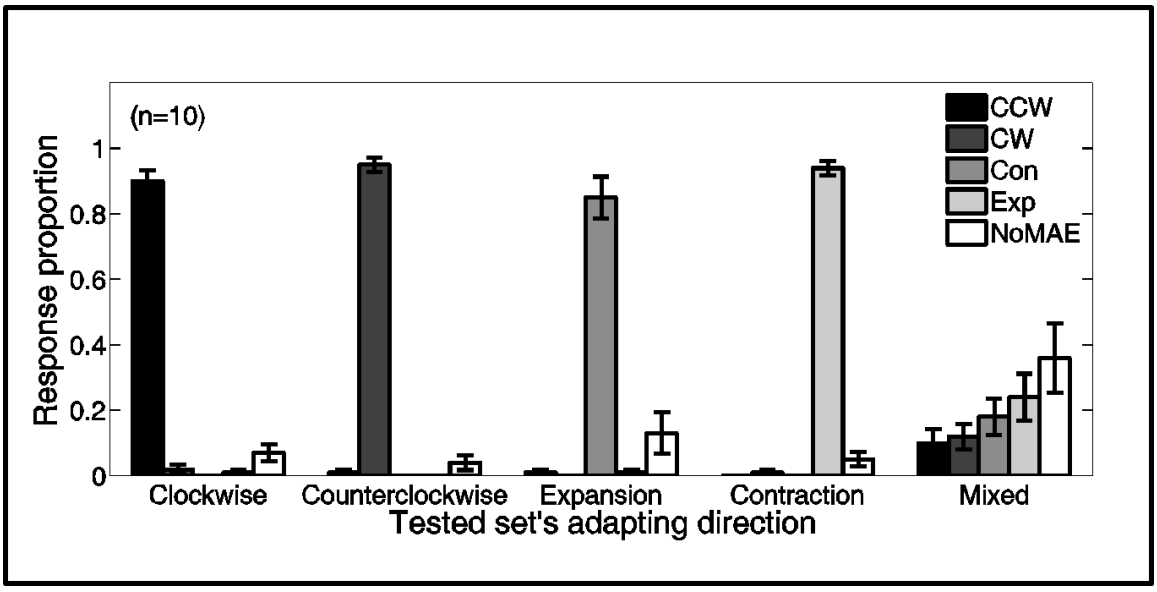

Figure 4.4. Results of Experiment 2. Bar groups indicate conditions. Different shades indicate reported MAE directions. Each error bar represents ±1 S.E.M.

Figure 4.4 shows the results of Experiment 2. For the four Single conditions (the four groups on the left of Figure 4.4), there is a clear trend that perceived MAE direction was opposite to the tested set's adapting direction. This result is qualitatively similar to the results in Experiment 1. Response proportions of this opposite MAE direction ranged from 85% to 95% for all four conditions, with each condition having a small proportion (4% to 13%) of trials in which the observer reported no MAE direction. On the other hand, for the Mixed condition, the most frequently chosen option was NoMAE (36%), and the directional responses were more evenly distributed (from 10% to 24%), suggesting that observers did not consistently perceive any specific MAE directions in this condition.

Two-stage Model

Previous electrophysiological studies found that adaptation changes the tuning properties of V1 neurons in the primate brain (Dragoi, et al., 2000; Kohn & Movshon, 2004). At the low level of processing, neurons appear to undergo some characteristic changes in their tuning properties after adaptation (Kohn, 2007; Schwartz, et al., 2007). A set of adaptation-altered tuning curves produces a pattern of population responses different from population responses without adaptation (Schwartz, et al., 2007). These altered population response patterns can produce the repulsive perceptual bias induced by motion adaptation (Schwartz, et al., 2007; A. A. Stocker & Simoncelli, 2009). It is possible that these local, adaptation-induced perceptual biases can be integrated by global-motion mechanisms, presumably at the MT level, to produce a global MAE. Such integration may explain various MAE phenomena after adapting to bidirectional

transparent motion patterns (Vidnyanszky, et al., 2002), including the integrated MAE (Snowden & Verstraten, 1999) and the orthogonal MAE (Grunewald & Lankheet, 1996). This section presents a two-stage model that computes local illusory signals and integrates them according to test locations. The proposed model provides a computational account to the MAE perceived after adapting to imperceptible motion stimulus in our study.

Figure 4.5 shows a schematic illustration of our model, which consists of two stages: 1) computation of local illusory signals at each location based on altered tuning curves after adaptation, and 2) integration of local illusory signals across test locations to produce the perceived MAE. In the first stage, velocity-selective motion detectors at each location changed their tuning properties. Such changes were induced by the local adapting velocity. Then, a local illusory signal was computed at each location, based on the population response pattern of the motion detectors when stimulated by a static stimulus. In the second stage, local illusory signals at the tested locations were integrated to produce perceived MAE for different test conditions.

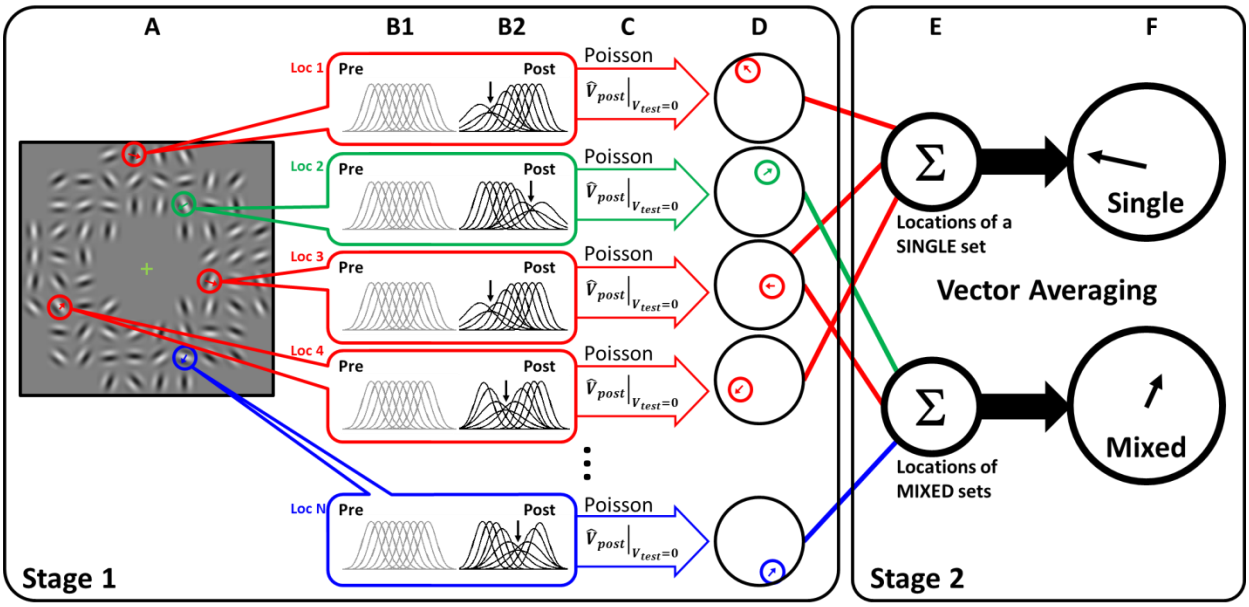

Figure 4.5. Schematic illustration of the model. Colors represent different sets of elements in the stimulus. Parameters (e.g., number of elements, velocity dimensions, etc.) are modified for illustration purpose. In Stage 1, the model computes an illusory velocity for each location as follows: In each trial, the model takes the drifting velocity at each location from the stimulus as input (A). At each location, different velocity-tuned motion detectors change their tuning properties based on the adapting velocity at that location (from B1 to B2). Based on the post-adaptation response pattern, the model computes an estimate on a static stimulus (C) and gives this "illusory" velocity as output from Stage 1 for each location (D). In Stage 2, the model combines local illusory velocities based on the test condition (E, Top: Single, bottom: Mixed), and gives one global MAE velocity (F) as the response for each trial. Stimulus and procedure parameters for the model matched exactly with those used in the psychophysical experiment.

Stage 1: Computation of local illusory motion signals

Inputs to the model were the adapting velocity $\boldsymbol{V_i^A} = \left(v_{x,i}^A, v_{y,i}^A\right)$ (where $\boldsymbol{V^A}$ indicates adapting velocity) and location (x_i, y_i) of each element i, for $i = 1, \ldots, 264$ (total number of elements), and the set of tested elements T for each trial, which is a subset of all 264 elements. At each location, there were 256 local motion detectors, with each detector j preferring a specific velocity $\boldsymbol{U_j} = \left(u_{x,j}, u_{y,j}\right)$. Preferred velocities were chosen from 64 directions (evenly-spaced around 360°) across 4 speed levels (evenly-spaced between 0.2 and 1 speed units). Pre-adaptation velocity tuning for each detector j was modeled using a bivariate Gaussian function

$$f_j\left(\boldsymbol{V};\ \boldsymbol{U_j}, \sigma_j^2\right) \propto \exp\left(-\left(\boldsymbol{V} - \boldsymbol{U_j}\right)'\boldsymbol{\Sigma^{-1}}\left(\boldsymbol{V} - \boldsymbol{U_j}\right)/\boldsymbol{2}\right),$$

,where

$$\boldsymbol{\Sigma} = \begin{pmatrix} \sigma_j^2 & 0 \\ 0 & \sigma_j^2 \end{pmatrix}$$

is the covariance matrix. Because we assume that velocities in the horizontal (x) and vertical (y) dimensions do not correlate and have the same variance, we can simplify the bivariate Gaussian:

$$f_j\left(\boldsymbol{V};\ \boldsymbol{U_j}, \sigma_j^2\right) \propto \exp\left(\frac{-\|\boldsymbol{V}-\boldsymbol{U_j}\|^2}{2\sigma_j^2}\right),$$

, where $\boldsymbol{V}$ is the stimulus velocity space, $\boldsymbol{U_j}$ the preferred velocity of the detector, and $\sigma_j^2 = 0.1$ the tuning width (constant for all detectors before adaptation). This tuning "surface" across the two-dimensional stimulus velocity space represents the mean activation or expected "firing rate" of each detector when stimulated with different velocities. Each tuning function was normalized so that $f_{min} = 5$ and $f_{max} = 22$, estimated based on previous electrophysiological findings (Bair & Movshon, 2004; Sceniak, Ringach, Hawken, & Shapley, 1999).

In each trial, tuning characteristics of local motion detectors were altered by the adapting velocity $\boldsymbol{V}_{\boldsymbol{i}}^{\mathrm{A}}$ at each location . We implemented three types of adaptation-induced changes in local tuning properties (Schwartz, et al., 2007), including repulsive shift of preferred velocity away from the adaptor, broadening of tuning width, and reduction of responsivity (Figure 4.6).

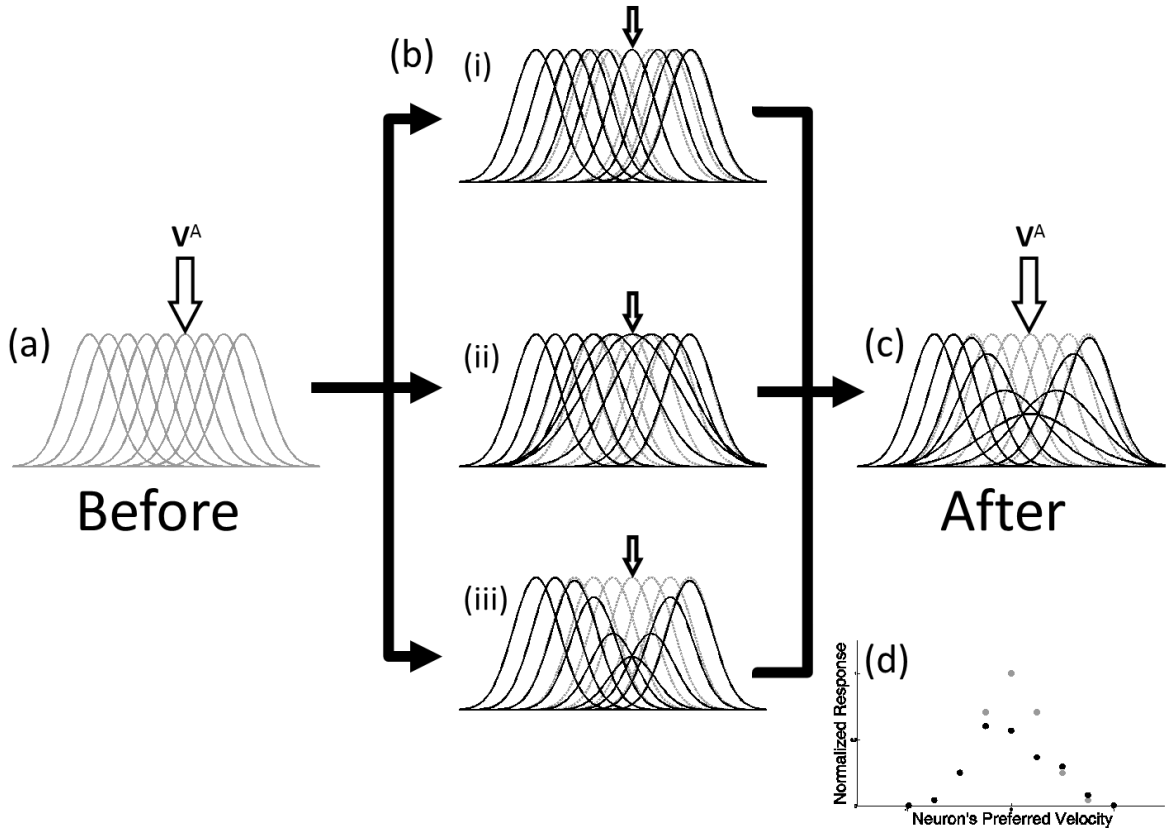

Figure 4.6. Illustration of adaptation-induced changes in local velocity tuning. Light and dark colors represent neural responses before and after adaptation, respectively. The illustration has been simplified to a one-dimensional velocity representation, while velocity had two dimensions in our implementation. (a) Before adaptation, local velocity detectors have homogeneous tuning properties. (b) Adapting to a velocity V^A leads to three types of changes in tuning properties of local detectors: (i) repulsive shift of preferred velocity away from V^A, (ii) broadened tuning width, and (iii) reduced responsivity. (c) These changes are combined to produce the altered velocity tuning functions for the motion detectors after adaptation. (d) The population response pattern on a static stimulus changes after adaptation, and produces the perceptual motion aftereffect.

First, the preferred velocity $\boldsymbol{U_j}$ of detector j after being adapted to velocity $\boldsymbol{V_i^A}$ at location i would be updated (Figure 4.6 (a)):

$$\boldsymbol{U_j}' = \boldsymbol{U_j} + \delta_{ij}\left(\boldsymbol{U_j} - \boldsymbol{V_i^A}\right)$$

, where

$$\delta_{ij} \propto \exp\left(\frac{-\left\|\boldsymbol{U_j} - \boldsymbol{V_i^A}\right\|^2}{2\sigma_\delta^2}\right)$$

In other words, δ_{ij}exaggerates the difference between the preferred velocity and the adapting velocity, to simulate the repulsive shift in preference caused by flank adaptation (Dragoi, et al., 2000). δ_{ij} was normalized so that $\delta \in [1, 1.25]$.

Second, tuning width of detector j (represented as the variance σ_j^2) was updated (Figure 4.6 (b)) as follows:

$$\sigma'^2_j = w_{ij}\sigma_j^2$$

, where

$$w_{ij} \propto \exp\left(\frac{-\left\|\boldsymbol{U_j} - \boldsymbol{V_i^A}\right\|^2}{2\sigma_w^2}\right)$$

, in which w_{ij} induces greater increase in tuning width for detectors when preferred velocity is closer to the adapting velocity, which reflects findings reported in previous electrophysiological studies (Dragoi, et al., 2000). Normalization operation was needed to limit w_{ij} within the range of [1, 1.2]. Finally, to implement the reduction in activation of detector j when presented with stimulus $\boldsymbol{V}$ (Figure 4.6 (c)), we scaled $f_j(\boldsymbol{V})$ by a factor, so that the post-adaptation activation $g_j(\boldsymbol{V})$ is given by

$$g_j(\boldsymbol{V}) = (1 - r_{ij})\, f_j(\boldsymbol{V})$$

, where

$$r_{ij} \propto \exp\left(\frac{-\left\|\boldsymbol{U_j} - \boldsymbol{V_i^A}\right\|^2}{2\sigma_r^2}\right)$$

r_{ij}, limited within the range of [0, 0.5], controls how much response suppression is applied to a detector's activation: The closer the detector's preference is to the adaptor, the greater the suppression, which is based on previous physiological findings (Dragoi, et al., 2000; Kohn, 2007). The three variances $\sigma_\delta^2, \sigma_w^2,$ and σ_r^2 control, respectively, the range of effects that adaptation has over the detectors for repulsive shift of preference, broadening of tuning width, and reduction of responsivity. In principle, these parameters could be independently manipulated to reflect different ranges of effects for different types of adaptation-induced changes. However, for simplicity, we kept all of them constant in our implementation, so that

$$\sigma_\delta^2 = \sigma_w^2 = \sigma_r^2 = 0.2$$

In summary, the post-adaptation velocity tuning of detector *j* after adapting to velocity $\boldsymbol{V_i^A}$ at location i is computed by

$$g_j\left(\boldsymbol{V};\ \boldsymbol{U_j}, \sigma_j^2, \boldsymbol{V_i^A}\right) = (1 - r_{ij})\ f_j\left(\boldsymbol{V};\ \boldsymbol{U_j}', \sigma'^2_j\right)$$

After implementing adaptation-induced turning changes of local motion detectors, we randomly sampled a simulated "firing rate" n_{ij} for a test velocity $\boldsymbol{V_{test}}$, which was taken to be the 1D drifting velocity of the Gabor element at location i. Firing rates were independently sampled from the Poisson distribution for each detector *j* at each location *i*, based on the post-adaptation tuning function $g_j(\boldsymbol{V_{test}})$:

$$n_{ij} \sim \text{Poisson}\left(g_j(\boldsymbol{V_{test}})\right)$$

Since static test stimuli was used in both of our experiments, $\boldsymbol{V_{test}}$ always equaled $\mathbf{0}$ in our implementation. Note that the model is capable of handling nonzero test velocities (i.e., testing for dynamic MAE).

We assumed that the local illusory signal at each location i is simply the post-adaptation representation of local velocity at that location. In order to read out this representation of velocity, we employed the strategy proposed by Jazayeri and Movshon (2006). We computed the log-likelihood $\log L_i(\boldsymbol{V})$ for a range of possible stimulus velocities $\boldsymbol{V}$ based on the population response profile of all detectors at each location i

$$\log L_i(\boldsymbol{V}) = \sum_{j=1}^{J} n_{ij} \log g_j(\boldsymbol{V})$$

We then obtain the local illusory velocity $\widehat{\boldsymbol{v}}_{\boldsymbol{i}}$ for location i based on the maximum-likelihood estimate:

$$\widehat{\boldsymbol{v}}_{\boldsymbol{i}} = \arg\max_{\boldsymbol{v} \in \boldsymbol{V}} \log L_i(\boldsymbol{V})$$

Stage 2: Integrating local illusory signals

In each experimental trial, the model gave a global MAE directional response, which was computed by integrating local illusory motion signals over the test locations selected for that trial. The general approach we used for spatial integration can be understood as a template-matching strategy, with the best-matched motion template being chosen as the output. This is similar to applying a winner-take-all read-out strategy on a large set of global motion detectors that are tuned for different global motion flows. Formally, let $\{\widehat{\boldsymbol{v}}\} = \{\hat{v}_{x,i}, \hat{v}_{y,i}\}$ be the set of local illusory motion vectors computed at all locations in Stage 1. We define an error function $E(\{\boldsymbol{v}\})$ for each global flow field template $\{\boldsymbol{v}\}$

$$E(\{\boldsymbol{v}\}) = \|\{\widehat{\boldsymbol{v}}\} - \{\boldsymbol{v}\}\|^2 = \sum_{i \in \boldsymbol{T}} \|\widehat{\boldsymbol{v}}_i - \boldsymbol{v}_{\boldsymbol{i}}\|^2,$$

where T denotes the set of test locations for that trial (given to the model as input), and $\{\boldsymbol{v}\}$ the global flow field template. $E(\{\boldsymbol{v}\})$ essentially represents the difference between the local illusory motion vectors $\{\widehat{\boldsymbol{v}}\}$ and the global flow field template $\{\boldsymbol{v}\}$. The estimated global MAE flow field is taken to be the global flow field template $\{\boldsymbol{v}^*\}$ that minimizes the error E

$$\{\boldsymbol{v}^*\} = \arg\min_{\boldsymbol{v}\in\boldsymbol{V}} E = \arg\min_{\boldsymbol{v}\in\boldsymbol{V}} \sum_{i\in\boldsymbol{T}} \|\widehat{\boldsymbol{v}}_i - \boldsymbol{v}_i\|^2$$

Then, by solving the equation

$$\nabla E = 0$$

, we can obtain a least-square estimate of $\boldsymbol{v}^*$.

For Experiment 1, we assumed the global MAE flow field templates to be translational and homogeneous across locations, so that $\boldsymbol{v_i} = \boldsymbol{v} = \{v_x, v_y\}$, for all locations i. The solution can then be derived analytically:

$$\boldsymbol{v}^* = \frac{\mathbf{1}}{\|\boldsymbol{T}\|} \sum_{i\in\boldsymbol{T}} \widehat{\boldsymbol{v}}_i$$

, which is identical to the solution given by a vector-average algorithm (Wilson, Ferrera, & Yo, 1992). An angular response was then obtained by the two-argument arctangent function $atan2(v_y{}^*, v_x{}^*)$, whose output range spans the whole range of 360°.

For Experiment 2, we defined our complex flow field templates based on the work by Wu and colleagues (2009), which assumes ideal, rigid complex flow fields for circular and radial motion templates. For circular flows, we assumed a constant angular velocity ω across locations:

$$\begin{cases} v_{x_i} = -\omega r_{y_i} \\ v_{y_i} = \omega r_{x_i} \end{cases};$$

and for radial flows, we assumed a constant expansion velocity e across locations:

$$\begin{cases} v_{x_i} = er_{x_i} \\ v_{y_i} = er_{y_i}, \end{cases}$$

where $r_{x_i}$and $r_{y_i}$are the horizontal and vertical distances, respectively, from location i to the center of rotation/expansion. Note that the signs of ω and e, respectively, determine the directions of rotation (clockwise vs counterclockwise) and radial motion (contraction vs expansion). Substituting these back into the error function $E(\{\boldsymbol{v}\})$ and applying the optimization algorithm, the least-square estimates for ω and e are, respectively,

$$\omega^* = \frac{\sum_{i\in \boldsymbol{T}}\left(\hat{v}_{y_i} r_{x_i} - \hat{v}_{x_i} r_{y_i}\right)}{\sum_{i\in \boldsymbol{T}} \|\boldsymbol{r}_i\|^2}, \; e^* = \frac{\sum_{i\in \boldsymbol{T}}\left(\hat{v}_{x_i} r_{x_i} + \hat{v}_{y_i} r_{y_i}\right)}{\sum_{i\in \boldsymbol{T}} \|\boldsymbol{r}_i\|^2},$$

where $\|\boldsymbol{r}_i\|^2 = r_{x_i}^2 + r_{y_i}^2$.

The model's response in the 4AFC task (Clockwise, Counterclockwise, Inward or Outward) was determined based on the magnitudes (speed) and signs (direction) of the estimates of the complex motion flow:

$$\begin{cases} if\ |\omega^*| > |e^*|, and \begin{cases} if\ \omega^* > 0, \text{Response} = \text{Counterclockwise} \\ if\ \omega^* \leq 0, \text{Response} = \text{Clockwise} \end{cases} \\ if\ |\omega^*| \leq |e^*|, and \begin{cases} if\ e^* > 0, \text{Response} = \text{Outward} \\ if\ e^* \leq 0, \text{Response} = \text{Inward} \end{cases} \end{cases}$$

Note that model's response in each trial could only be one of the four directional responses. Unlike our Experiment 2 with humans, the model would not produce a "No-MAE" response.

Model Predictions

Distributions of the model's predicted MAE directions in the two conditions are shown in Figure 4.7A, with human responses in Experiment 1 replotted for comparison. For the Single condition, the central tendency of the model's predicted MAE directions (Circular mean angle = 180.63°) was found to align closely with that of human's reported MAE directions (Circular

mean angle = 178.34°). However, the two distributions differed greatly in variability, with human responses having a much wider spread (Circular standard deviation = 65.18°) than model predictions (Circular standard deviation = 15.75°). For the Mixed condition, distribution of model's responses was not different from a uniform distribution (Hodges-Ajne test, $p = .934$). This shows that our model can predict the overall pattern of MAE direction perceived by humans in both conditions.

We further compared model's prediction with human's perception by computing the prediction error in each trial, which was taken to be the angular difference in reported MAE directions between human and model. If the model's predicted MAE direction perfectly matched with that reported by human observer, the prediction error for that trial would be 0°. Figure 4.7B shows the polar histograms of prediction errors for both conditions. In about 60% of the trials in the Single condition, prediction error was within 45°. This suggests that the model, despite its simple vector-averaging integration stage, can reliably predict perceived MAE directions in the Single condition. For the Mixed condition, the distribution of prediction errors appears to be fairly uniform (Hodges-Ajne test, $p = .11$), with a weak bimodal trend peaking at 0° and 180°. In about 25% of the trials, prediction error was within 45° of the MAE direction perceived by humans. However, in about the same number of trials (27.5%), the prediction error was within 45° of the opposite angle to the MAE direction reported by humans.

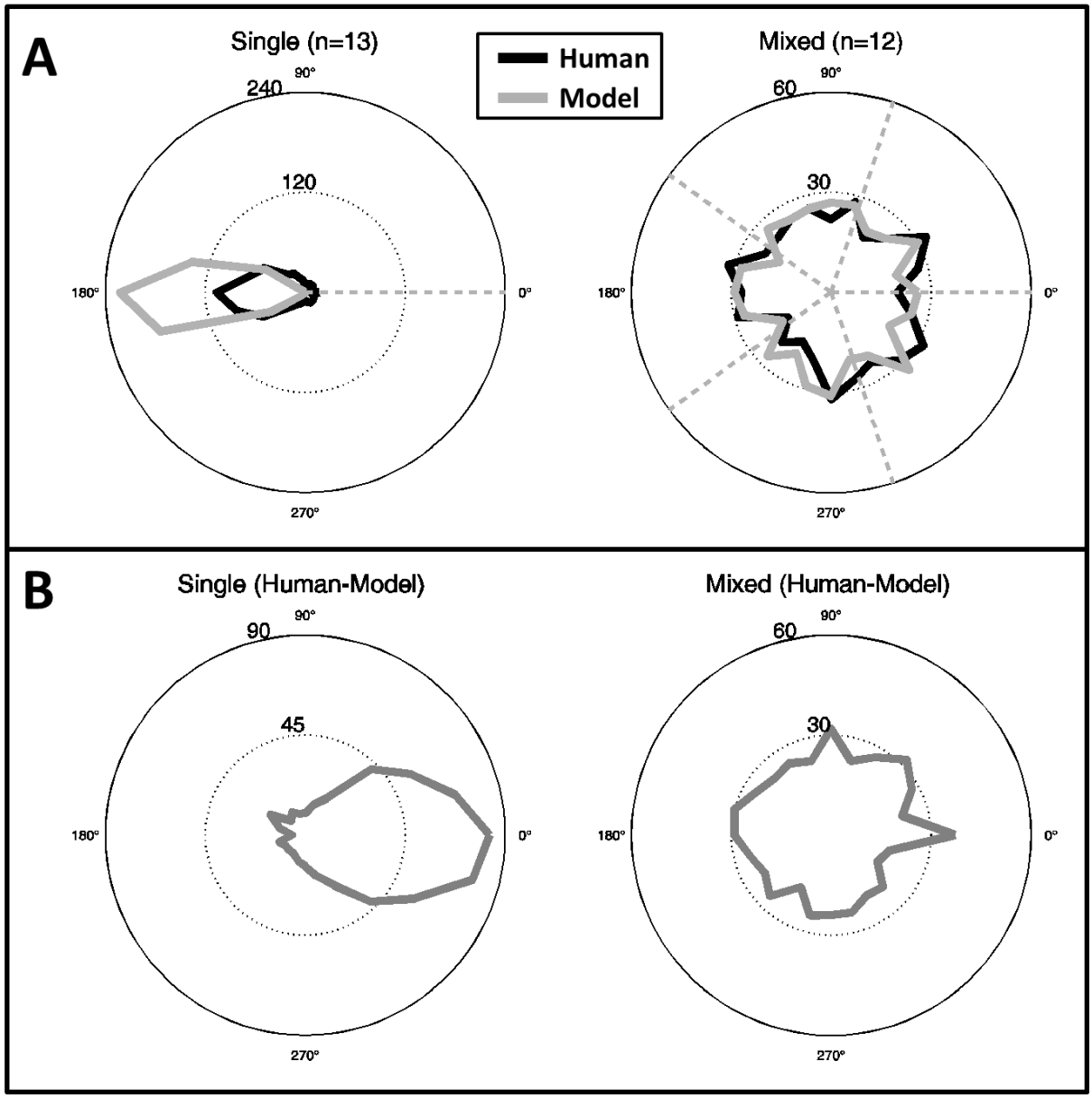

Figure 4.7. Results from model's prediction, compared with human data obtained in Experiment 1. (A) Distributions of MAE directions predicted by the model (light solid line) for the two conditions. Human data (same as those plotted in Figure 4.3) were replotted (dark solid line) for comparison. Note that scales for the two conditions are made different for better illustration. (B) Distributions of prediction error, which was computed as the difference in responses between human observers' and model's predictions for the two conditions.

Figure 4.8 shows the results of model's prediction for Experiment 2. The "No MAE" option was not listed among the alternatives for the model because it was forced to give one of the four directional responses based on the least-square method. In the four Single conditions (i.e., the four groups on the left), the model always picks the MAE direction that is opposite to the adapting direction. In other words, the complex MAE direction generated based on integrating local illusory motion signals is invariably opposite to the global adapting direction. This suggests that the small variability observed in humans might be due to some other sources, which could be noise at higher levels of processing or noise from other modules (e.g., motor noise in response). Overall, model predictions can qualitatively account for human performance. In the Mixed condition, the model gives a distribution of responses that appears to be uniform, different from the slight dominance of expansion MAE found in human results that may be due to response bias.

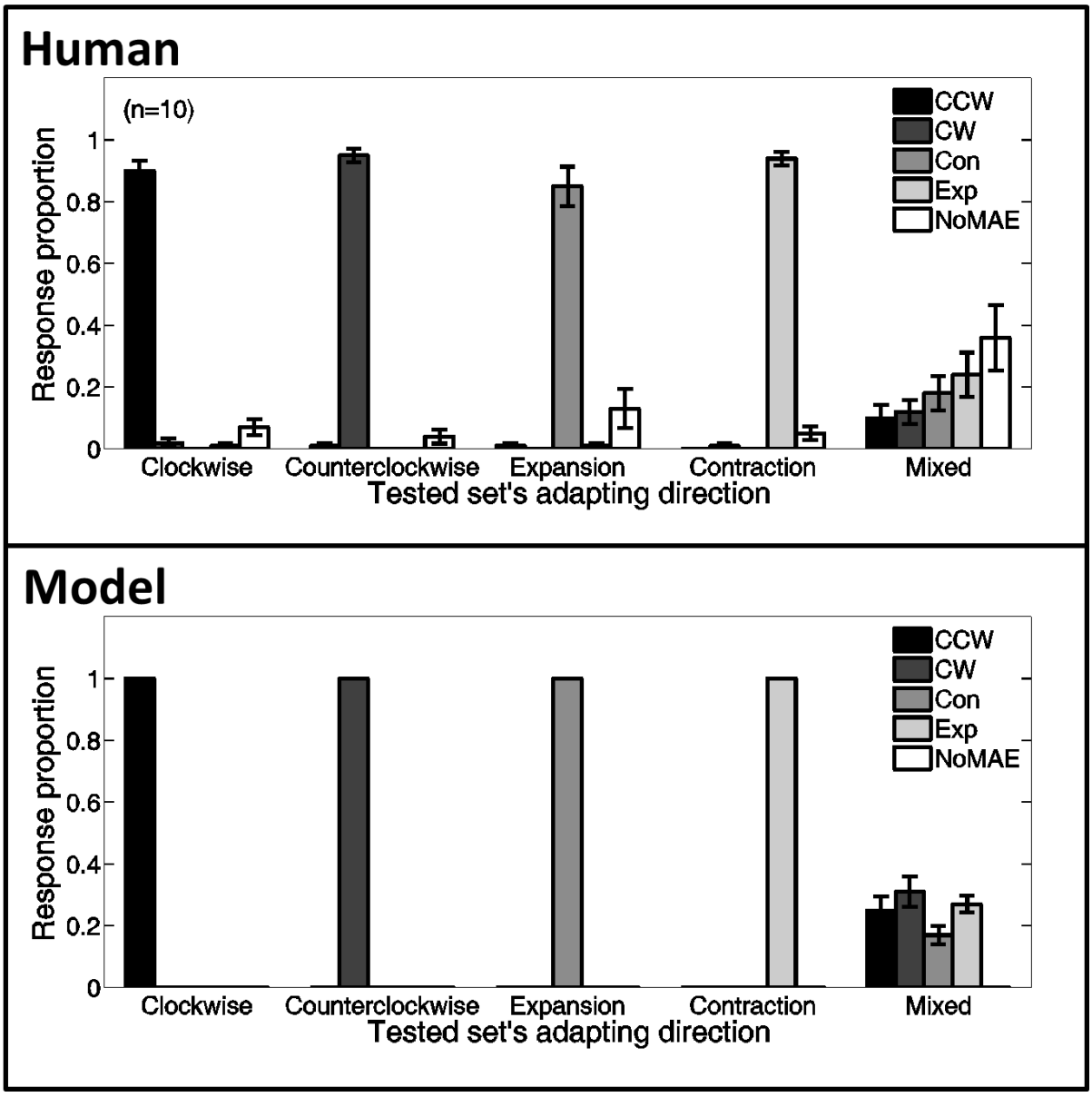

Figure 4.8. A comparison between human results (top, same as Figure 4.4) and model predictions (bottom) for Experiment 2.

General Discussion

We found that adaptation of local motion detectors *per se* is sufficient to produce a global directional MAE (translational, rotational or radial). This MAE generation process did not require the conscious perception of any of the adapting directions. Furthermore, we found that a two-stage model that computes adaptation-induced changes of local tuning properties and integrates local illusory signals can explain the perceived global MAE directions. Taken together, these results support the existence of a mechanism in the visual system that integrates local illusory signals to produce a global MAE percept.

There are two interesting points about the comparison between our psychophysical results and model prediction. First, in the Single condition in Experiment 1, while our model's averaged predicted MAE direction aligns with the MAE direction perceived by humans, the distribution of predicted directional responses was found to be narrower than the distribution of responses made by humans. Indeed, we tried a range of parameter sets and found similar results. If the visual system produced MAE percepts in exactly the same way as our model does, the distribution of human responses would have been much narrower than what we found in the psychophysical experiment. As the only source of noise in our model comes from the assumption that activation of local motion detectors is a Poisson process, our results imply that this variability at the local level cannot fully account for variability in perceived MAE direction. We speculate that the additional variability in human responses found in our experiment might be due to noise in other motion processing stages (e.g., noise in the motion-integration process). It is also possible that noise in other modalities (e.g., motor noise in making responses) may contribute to the variability in human responses. Second, in the Mixed condition in both experiments, our model does not predict human perception as well as it does in the Single

condition. It may be due to the fact that our model employs a simple, template-matching integration strategy. It has been shown that a Bayesian motion-integration strategy with a suitable functional form for the prior (H. Lu, Lin, Lee, Vese, & Yuille, 2010) can produce human-like responses in a 2AFC task across a wide range of motion coherence. Therefore, it is possible that a more sophisticated motion-integration strategy at Stage 2 may produce predictions that are more human-like. It should be noted that our model does not tie to any particular strategy of motion integration. In fact, in the present study, our goal is to propose a general framework for how adaptation-induced neural effects can be related to perceptual aftereffects. The formulation of our model here is only one of the many ways in which such propagation could be implemented.

Our empirical and computational findings provide direct evidence for the theory proposed by Vidnyanszky and colleagues (2002), who proposed that some apparently intriguing MAE phenomena are, in fact, due to the integration of local illusory signals. This theory has been reinforced by later physiological studies in motion adaptation (Kohn & Movshon, 2003, 2004), and a psychophysical study in face adaptation (Xu, et al., 2008), which show that adaptation effects at low level of processing can be propagated downstream to influence later stages of processing. This downstream propagation may be one of the sources of the perceived, high-level aftereffects. In the present study, our psychophysical and computational results clearly demonstrate how such propagation of low-level, local illusory signals is accomplished via spatial integration of local illusory signals in the case of motion adaptation. This finding connects adaptation-induced neural changes to perceived MAE and provides a unified description of how neural adaptation might lead to perceptual aftereffects.

Few previous studies have specifically investigated the conditions under which a global MAE is perceived with minimal, if any, global motion adaptation. In the present study, we attempted to suppress global motion adaptation using a multidirectional adaptor. One may consider this technique as "flooding" the direction-selective global mechanisms in order to minimize directional adaptation at the global-motion processing level. Combining with the multiple-aperture stimulus (Amano et al., 2009), we believe this method makes it possible to cleanly isolate different processing levels, in terms of evaluating their contributions in producing post-adaptation perceptual aftereffects.

It should be noted that the present study focuses on the mechanism that integrates local illusory signals and generates an aftereffect percept at the global level. There are other mechanisms that can produce an aftereffect percept, possibly without involving the integration mechanism. Indeed, many previous motion adaptation studies found a global MAE with minimal or no local adaptation, e.g., using RDKs with limited dot lifetime (e.g., Grunewald & Lankheet, 1996; F. A. J. Verstraten, et al., 1994) or testing the MAE at non-adapted locations (e.g., Lee & Lu, 2012; Scarfe & Johnston, 2011; Snowden & Milne, 1997), etc. These findings clearly demonstrate that the adaptation at the *global*-motion processing level *per se* is also a mechanism to produce a global MAE percept.

It appears that these two mechanisms, namely, 1) global motion adaptation and 2) integration of local illusory signals, can independently produce an aftereffect percept with minimal contribution from the other. Nonetheless, how do the two mechanisms interact to produce and aftereffect percept? How may one mechanism affect the other when both are operating? It has been shown that high-level adaptation may have some top-down influence on low-level adaptation processes (Scarfe & Johnston, 2011), and may dominate the aftereffect

percept when low-level adaptation was minimal (Lee & Lu, 2012). It will be interesting to investigate the relationship between the two aftereffect-generation mechanisms in the hierarchy of motion processing, and the conditions under which the two mechanisms may simultaneously contribute to the perception of aftereffects.

Our findings may also shed lights on the functional goal of sensory adaptation in general. Most previous studies used perceptible global motion directions as the adapting stimulus. Because, at some level, the system must "know" the motion directions we are presented with (e.g., leftward, expansion, clockwise rotation, etc.), it is reasonable for the system to be adaptive to these perceptible directions. However, in the present study, motion directions were imperceptible in the adapting environment. Theoretically, the system could have avoided being adaptive to such apparently random motion, because it does not represent a meaningful motion environment. Nonetheless, our findings suggest that our visual system remains adaptive to the local motion signals, to the extent that adaptation-induced effects across locations can be subsequently integrated to generate a global MAE. In other words, when no specific global adapting directions are available, the visual system can retain local motion information, which is inherently ambiguous in signaling the global motion direction (Marr & Ullman, 1981). Such retention of local information is so precise that the perceived global aftereffect direction can be determined by test locations. It has been proposed that sensory adaptation can be understood as a process by which the system, in response to changes in the environment, recalibrates its response characteristics and enhances its coding efficiency (Barlow, 1990; C. W. Clifford, Wenderoth, & Spehar, 2000). From a hierarchical-processing perspective, our findings suggest that the system can enhance its coding efficiency not only by recalibrating its

responses characteristics of high-level processing units, but also by temporarily altering response characteristics of low-level detectors. It is possible that integration of the low-level readjustment effects by some high-level mechanism is a means for the hierarchical system to improve its overall coding efficiency. It would be useful for future computational studies to investigate how this bottom-up, local-to-global, adaptation strategy may lead to more efficient coding in the system.

Chapter 5

Measuring Repulsive Direction Aftereffect:

Local vs Global Adaptation

Introduction

Adaptation can induce changes at different levels of visual processing. The previous two chapters systematically examine motion aftereffect, and trace the effects of adaptation through the motion processing hierarchy. In this chapter, we will investigate the other behavioral hallmark of sensory adaptation, the perceptual repulsion (PR) effect (C. W. Clifford, et al., 2007). Adapting to a particular direction causes the perceived direction of test motions to be shifted away from the adapting direction. In particular, after adapting to a certain sensory stimulus (e.g., upward motion direction), the subjective percept of a subsequently presented test stimulus (e.g., 45° clockwise from upward) is biased *away* from the adapting stimulus (e.g., percept = 60° clockwise from upward). Such adaptation-induced repulsion effect has been widely observed in many perceptual domains, including orientations (the tilt aftereffect, Gibson & Radner, 1937), motion (Levinson & Sekuler, 1976; Schrater & Simoncelli, 1998; A. A. Stocker & Simoncelli, 2009), faces (Webster, Kaping, Mizokami, & Duhamel, 2004) and biological motion (Troje, Sadr, Geyer, & Nakayama, 2006).

In the perception of motion direction, the direction aftereffect (DAE) is an important measure to quantify the strength of adaptation-induced repulsive effect. It is measured as the angular difference between the direction of the true testing stimulus and perceived direction. Curran and colleagues (Curran, et al., 2006) investigated which level of adaptation mainly drives the DAE.

They first measured how DAE strength (in terms of degrees of angular bias) changed as a function of motion speed for each observer, so that they have a lookup table from stimulus speed to DAE strength for each individual observer. They found that almost all observers showed a speed tuning in DAE. Then, they made use of this speed-tuned properties to test whether local or global adaptation mainly drives. Then, when arbitrating between the contribution of local and global adaptation in producing DAE, they proposed a "local" model and a "global" model for generating DAE. The "local" model assumes the speed-tuning of DAE is largely due to the adaptation of local motion detectors. This model predicts the strength of DAE by applying the speed-tuned function on each stimulus dot to obtain a set of local "DAE"s, then averages them to obtain one DAE measure. The "global" model assumes the observed speed-tuning is caused by high-level adaptation, so that a DAE measure can be obtained by first averaging the speeds of all the dots, then readout a DAE from that averaged speed. The authors found that the variation in DAE strength was more consistent with a local-motion model than with a global-motion model. They suggested that the DAE is mainly driven by the adaptation of detectors at the local-processing stage in the motion system. However, this finding is inconsistent with the tuning changes found in MT (Kohn & Movshon, 2004), and cannot be reconciled with the global-adaptation account for the spatial contextual effect of direction repulsion (O. J. Braddick, Wishart, & Curran, 2002; Marshak & Sekuler, 1979; Mather & Moulden, 1980). Moreover, what Curran and colleague found may be better interpreted as that the speed-tuning property of DAE is based on local speed instead of global one. If one could independently induce adaptation at the low and the high levels of motion processing, it would be possible to evaluate the contribution of the adaptation of neural units at different levels in the process of generating perceptual repulsion effects.

The challenge is to design an appropriate stimulus that allows independent adaptation-induced changes to the stimulation of the low- and the high-level motion detectors. We will address this issue in this chapter. In the motion-processing literature, researchers have used different psychophysical stimuli or technique to adapt detectors or suppress their adaptation at specific levels of processing. For low-level detectors, drifting gratings have been widely used as an inducer for adaptation (Anstis, et al., 1998; Kohn & Movshon, 2003, 2004; Nishida & Ashida, 2001). Gratings allow researchers to specifically probe low-level, local motion detectors, because they allow researchers to independently manipulate various aspects of motion, including stimulus location, orientation, spatial frequency and temporal frequency, for which low-level motion detectors are known to be selectively tuned (Adelson & Bergen, 1985; Ibbotson & Price, 2001; van Santen & Sperling, 1985). To suppress low-level adaptation, some researchers have probed perceptual aftereffect at locations or regions at which adaptation stimulus has not been presented (Kohn & Movshon, 2003; Scarfe & Johnston, 2011; Snowden & Milne, 1997; Weisstein, et al., 1977). Because low-level detectors usually have small receptive fields, one could minimize local adaptation by adapting the motion system in certain parts of the visual field, and then test it at the non-adapted regions or locations. As high-level, global motion units usually integrate motion signals across relatively large receptive fields (Morrone, et al., 1995), the aftereffect observed at these non-adapted regions is generally assumed to be caused by the adaptation of these global detectors. Indeed, using this technique, researchers have found what has been known as the "Phantom" aftereffect (Lee & Lu, 2012; Price, et al., 2004; Snowden & Milne, 1997; Weisstein, et al., 1977). However, there has been no study attempting to suppress high-level, global adaptation while maintaining low-level, local adaptation, as what I proposed in the previous chapter.

For the adaptation of motion direction, one possible way to minimize the adaptation of direction-selective units at the global level of motion processing is to embed multiple motion directions within a single motion pattern. Population response of neurons in MT depends on the underlying composition of motion directions in a stimulus pattern (Treue, et al., 2000). While there only two widely-separated directions are presented, population response is a bimodal distribution, which possibly gives rise to the percept of transparent motion (Qian & Andersen, 1994; Treue, et al., 2000). However, since there is a limit of bandwidth in their direction tuning (Bair & Movshon, 2004), when more and more directions were embedded within the same spatial region in the visual field, the population response of direction-selective units would become flat and incapable of encoding any specific directions. This is further supported by the psychophysical finding that humans are unable to perceive, for example, five evenly separated directions embedded within a single motion pattern (Greenwood & Edwards, 2006a, 2006b, 2009). Therefore, it is possible to "overload" the direction-selective global motion units with a multidirectional pattern in order to suppress direction-specific activation or adaptation at the global level of motion processing. By suppressing the contribution from the global level of motion processing, one could unveil how adaptation-induced effects originated from low-level, local motion processing are integrated to produce a global perceptual aftereffect.

No study so far has combined these techniques to investigate level-specific contribution in generating perceptual repulsion effect, and quantitatively measure the strength of the effect induced by adaptation at different processing levels. In this chapter, I will present empirical studies that address this issue using a novel adaptation paradigm. We combined the multiple-aperture stimulus (Amano, et al., 2009b), which allows independent manipulation of local and global motion signals, with the above-mentioned technique of "direction overloading" to study

adaptation in the motion-processing hierarchy. Specifically, we independently manipulated the adaptation states of the local and the global levels of motion processing, and measure the strength of PR in global motion direction (a.k.a. the direction aftereffect or DAE). In Experiment 1, using different psychophysical methods in measuring the DAE, we found that DAE was stronger when adaptation was specifically introduced at the global level than when it was at the local level. In Experiment 2, we found that the local-level DAE was orientation-specific, and was not strengthened even when PR in orientation (a.k.a. the tilt aftereffect or TAE) was introduced to couple with DAE. In Experiment 3, we found that the DAE arising from local adaptation was reduced when both the adapting and the test stimuli were locally unambiguous (as plaids).

General Methods

Design

To tease out the contribution of adaptation at the local and the global levels of motion processing in producing the DAE, we introduced two independent variables: First, test locations were manipulated to control the adaptation of the local level. Aftereffect was either tested at locations at which motion stimulus was presented during adaptation (Adapted) or at locations at which motion stimulus was absent during adaptation (Non-adapted). Second, the number of global adapting directions was manipulated to control the direction-specific adaptation at the global level. In the condition of all elements exhibited one coherent global motion direction during adaptation (1Dir), we expect strong global direction adaptation In the condition of five global directions introduced to the adapting stimulus (5Dir), we expect minimal direction-specific activation, and, thus, adaptation, at the global level, as evidenced in Chapter 4. Crossing these

two independent variables resulted in a 2-by-2 factorial design, which consisted of four conditions, named by the level(s) at which adaptation was introduced: BOTH (1Dir-Adapted), GLOBAL (1Dir-Non-adapted), LOCAL (5Dir-Adapted) and NEITHER (5Dir-Non-adapted) (right panel of Figure 5.1).

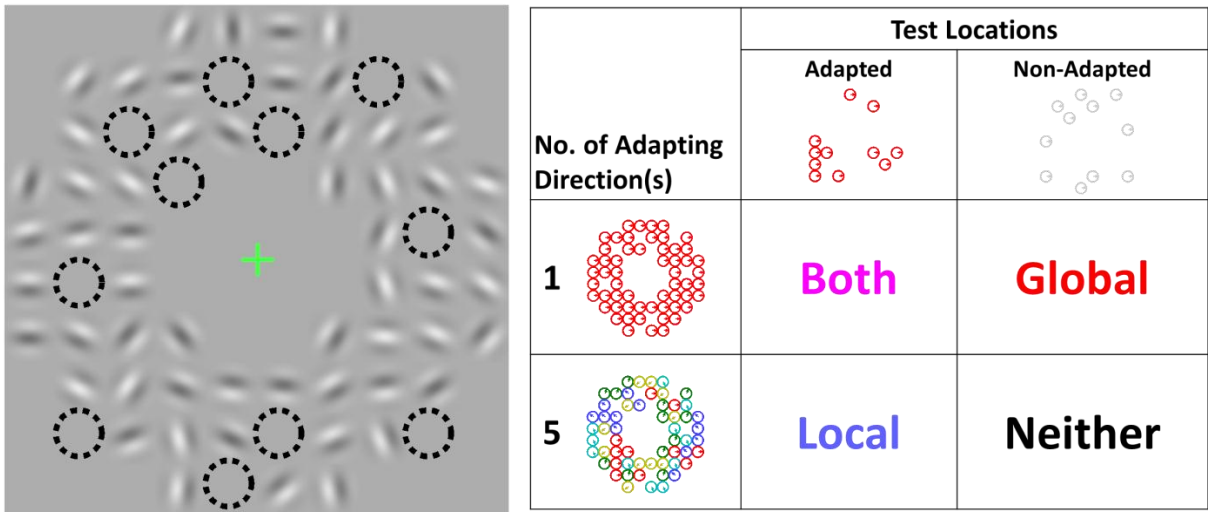

Figure 5.1. Stimulus and Design of the Experiments. Left: Illustration of the stimulus. Dashed circles (not shown in experiments) highlight the non-adapted locations in the stimulus. Dimensions have been halved for illustration purpose. The real stimulus contained about 4 times more elements. Right: The 2x2 design. In the cartoons, circles represent locations of elements. Colors represent the assignment of elements to different sets that were assigned different global adapting directions (indicated by the arrow within each circle). Elements with the same color were assigned the same global adapting direction.

Stimulus & Apparatus

The stimulus (left panel of Figure 5.1) consisted of 264 Gabor elements. Each element was a sine-wave grating, windowed by a stationary Gaussian function with a standard deviation of

0.21°. Spatial frequency was 2 cycles/° for all elements. Contrast was fixed at a low level of 0.05 (Michelson contrast) for all experiments, in order to enhance the perception of global motion based on spatial integration of local motion signals in the multiple-Gabor stimulus (Takeuchi, 1998). The diameter of each element spanned a visual angle of 1°. Positions of elements were fixed within a 20-by-20 grid, and arranged in a circular configuration, so that elements were presented within an annulus spanning 4°–10° of visual angle. Elements were tightly packed, so that separation between any two adjacent cells was zero. In other words, the centers of two adjacent Gabors were 1° apart. The whole pattern was centered within a larger circular aperture, with radius spanning 13° of visual angle. The background luminance value within this large aperture was at the mean (i.e., gray), while pixels outside this aperture were set to black. This circular aperture was to minimize possible directional bias in responses caused by the rectangular frame of the monitor.

The grating of each element was set to drifting motion, while its position remained constant. Orientation of each element θ was independently and uniformly sampled in the range of the 0°–180°. A global motion vector, with speed v and direction α, was assigned to each element. The local drifting speed u for each element was then computed based on the following equation:

$$u = v\sin(\alpha - \theta)$$

v was set to be 2°/s for all experiments. α depended on the number of adapting directions (see descriptions below).

The 264 elements were randomly split into six sets allowing to manipulate the number of global adapting directions and test locations in the experiments. We randomly chose one set to be the *reference* set, and another one the *non-adapted* set. During adaptation, regardless of conditions, all elements in the non-adapted set were hidden, so that there would not be any adapting motion

signals at those locations. In the adapting stimulus, the number of global adapting directions was manipulated by assigning different global motion directions to the five sets (i.e., reference and the other four). In the 1Dir conditions, the same global motion direction r was assigned to all elements, i.e., $\alpha = r$. In the 5Dir conditions, five global motion vectors, all with the same speed but different directions (equally spaced around 360°) were assigned to the five sets. Formally, the global direction a_i for each set i was

$$\alpha_i = r + (i - 1)(72°)$$

, with α_1 being the global direction for the reference set, so that $\alpha_1 = r$. In some experiments in which observers did a direction judgment task (see General Procedure), the reference direction r was randomly varied across blocks of trials, in order to avoid any direction-specific effects. In other experiments in which observers did a direction discrimination task, the reference direction r was always 45° away from the upward direction to introduce direction-specific adaptation effects (see General Procedure).

In the test stimulus, all elements were assigned the same global motion vector with global speed being the same as the adapting speed (i.e., 2°/s). Global direction depended on the task and experimental procedure (see General Procedure). Test locations were manipulated by choosing which set of elements to be displayed as a test stimulus. In the Adapted conditions, only elements in the reference set were displayed during the test, so that, at these test locations, adapting motion signals had all been consistent with one global adapting direction, i.e., r. In the Non-adapted conditions, only elements in the non-adapted set were presented for the test, so that the aftereffect probed at these locations should have minimal (if any) contribution from the adaptation of low-level, location-specific motion detectors.

For all experiments, MATLAB and PsychToolbox (Brainard, 1997; Pelli, 1997) were used to generate the stimuli, which were presented in a dim room on a Viewsonic CRT monitor (refresh rate = 75 Hz, resolution = 1024 pixels ×768 pixels). Viewing distance was kept constant at 57 cm using a chinrest and forehead rest, resulting in a visual angle of 2.01 arcmin for each pixel on the monitor. We used a Minolta CS-100 photometer to calibrate the monitor, and converted a luminance range of 0–146.5 cd/m^2 into a linear lookup table for 256 intensity levels. Participants were undergraduate students at the University of California, Los Angeles (UCLA), participating for course credit or for monetary reward. All participants had normal or corrected-to-normal vision and were naïve to the purpose of the experiments. The experiments were approved by UCLA's Office for Protection of Research Subjects.

General Procedure

In all experiments, observers were instructed to maintain fixation at a cross located at the center of the display. Because low-level motion detectors have relatively small receptive size, maintaining fixation would help stabilizing the effect of adaptation at induced specific retinotopic locations.

Each trial consisted of three phases: adaptation, test, and response. In order to inform observers of the different phases in each trial, the fixation cross was green, red, and white during the adaptation, test, and response phases, respectively. First, in the adaptation phase, observers viewed the adapting stimulus for a relative long duration (Initial adaptation: 45-60 seconds; top-up adaptation: 5-10 seconds; exact duration differed over experiments). After adaptation, observers were presented a blank screen (500ms) in the stimulus area, followed by the test stimulus that lasted for 720ms (54 frames). Last, in the response phase, all elements were removed and observer made a response, normally within 3 seconds. These durations were chosen

to match with those used in Schrater and Simoncelli's (1998) study, which found the characteristic repulsion curve for motion adaptation.

Observers performed one of the two tasks described below, depending on the specific experiment setup. The first task was a direction judgment task, in which a simulated dial was displayed during the response phase, with a white circular frame (radius = 11° of visual angle) and a short (1.5°) red indicator line extending outward from the circular frame to the edge of the large aperture. Observers were instructed to turn this simulated dial (i.e., move the red indicator line around the white circle) by pressing the left or the right arrow key to indicate their perceived global direction of motion during the test phase. When the red indicator was pointing at the desired direction, the observer pressed spacebar to submit this direction as a response for that particular trial. The initial direction to which the red indicator line pointed to was uniformly sampled around the 360° range in order to minimize any systematic bias caused by the default position. The second task was a direction discrimination task, in which nothing except the fixation cross was displayed on the monitor during the response phase. The task was to determine whether the global motion direction of the test stimulus was to the left (counterclockwise) or to the right (clockwise) relative to the upward direction.

Trials were grouped into blocks of 12–16 trials, with the first trial in the block having the longest adaptation duration. Top-up adapting motion patterns within the same block were random excerpts of the adapting stimulus of the first trial, so that the adapting stimulus of each trial within the same block showed exactly the same element configuration (e.g., orientations, assignment of sets, assignment of global adapting directions, etc.).

Experiment 1: Comparing DAE resulting from local and global adaptation

In Experiment 1, we aimed to investigate whether adaptation at the local and the global levels of motion processing would lead to different strengths of DAE. We employed the 2-by-2 factorial design described in General Methods, and instructed observers to perform the direction judgment task (See General Procedure under General Methods). In the Both condition, when adaptation was induced at both local and global levels, we expected DAE to be the strongest. In the Neither condition, when neither of the levels adapted, we expected zero DAE. Furthermore, if DAE depended on the level at which adaptation occurs, perceptual biases in the Global and the Local conditions would differ.

Experiment 1A: Assessing DAE across the full range of test directions

Methods

Stimulus was as described in the General Methods. The 2-by-2 factorial design was employed, resulting in the four conditions named by the level(s) at which adaptation was induced: Both, Global, Local and Neither. For each condition, 16 directions t_i, for $i = 1, 2, \dots, 16$, relative to the reference adapting direction r were tested. The 16 directions were evenly spaced within the range [–168.75°, 168.75°] with an interval of $22.5°$,

$$t_i = r + (-168.75° + 22.5°(i - 1))$$

Thirty-six observers participated in this experiment, with 9 in each of the four conditions. Observers performed the direction judgment task (see General Procedure). In order to assess whether there was any preexisting biases in direction judgment, observers first completed a Test-only session. In this session, the adaptation phase was removed, and observers viewed the test stimulus only and judged its motion direction. There were 5 blocks of 16 trials. The 16 test

directions t_i were randomly assigned to the 16 trials within a block, so that each directions was tested 5 times (once in each block) for each observer in each session. Afterwards, observers completed the Adapt-Test session, which was identical to the previous Test-only session, except that the adaptation phase preceded the test phase in each in each block (see General Procedure). In this Adapt-Test session, each block was assigned the reference adapting direction r, which was randomly sampled from a uniform distribution, with the constraint that the five r values were at least 60° apart in order to avoid massing all r values within a limit range for a particular observer. A 60-second rest period was given in between every two blocks to allow time for the previous block's adaptation effects to dissipate. Initial and top-up adaptation durations were 60 seconds and 10 seconds, respectively. Observers completed two short practices blocks at the beginning of each session to familiarize themselves with the task.

Results

Perceptual bias was calculated by the difference between the reported and the true directions. Therefore, a 0° bias indicates a reported direction that coincides with the true stimulus direction. In some trials, the absolute values of perceptual bias were as large as 90° or above. As these appeared to be outlier responses relative to the other trials, we removed these trials (234 out of 2880 trials, ~8%) in our analysis. The averaged perceptual bias for each observer in each session was then computed.

The upper panel of Figure 5.2 shows the perceptual bias for all four conditions, averaged across observers. Perceptual biases for both the Test-only and the Adapt-Test sessions were displayed. Perceptual bias was around zero in the Test-only session (thin-gray lines) across all conditions, suggesting that there was not systematic bias in direction judgment without motion adaptation. For the Adapt-Test session (thick, colored lines with open symbols), perceptual bias appears to

differ across conditions. Consistent with our expectation, the characteristic "S" shape of a repulsive bias curve was found in the Both condition but was absent in the Neither condition. A similar trend of DAE was found in the Global condition, but the bias curve for the Local condition does not show much DAE.

In order to quantify the strength of DAE, we fitted the average perceptual bias b_i for each observer using the following function

$$\hat{b}_i = a \sin t_i \exp(\cos t_i)$$

, where a indicates the size of repulsive bias. When a is positive, the fitted curve would yield the characteristic repulsive bias curve reported in the literature. When a is negative, the fitted bias curve would be in the opposite direction, i.e., perceptual attraction. The magnitude of a represents how strong the perceptual bias is, particularly around a test direction of $\pm 45°$. Therefore, fitted value of a can serve as the bias index for each observer.

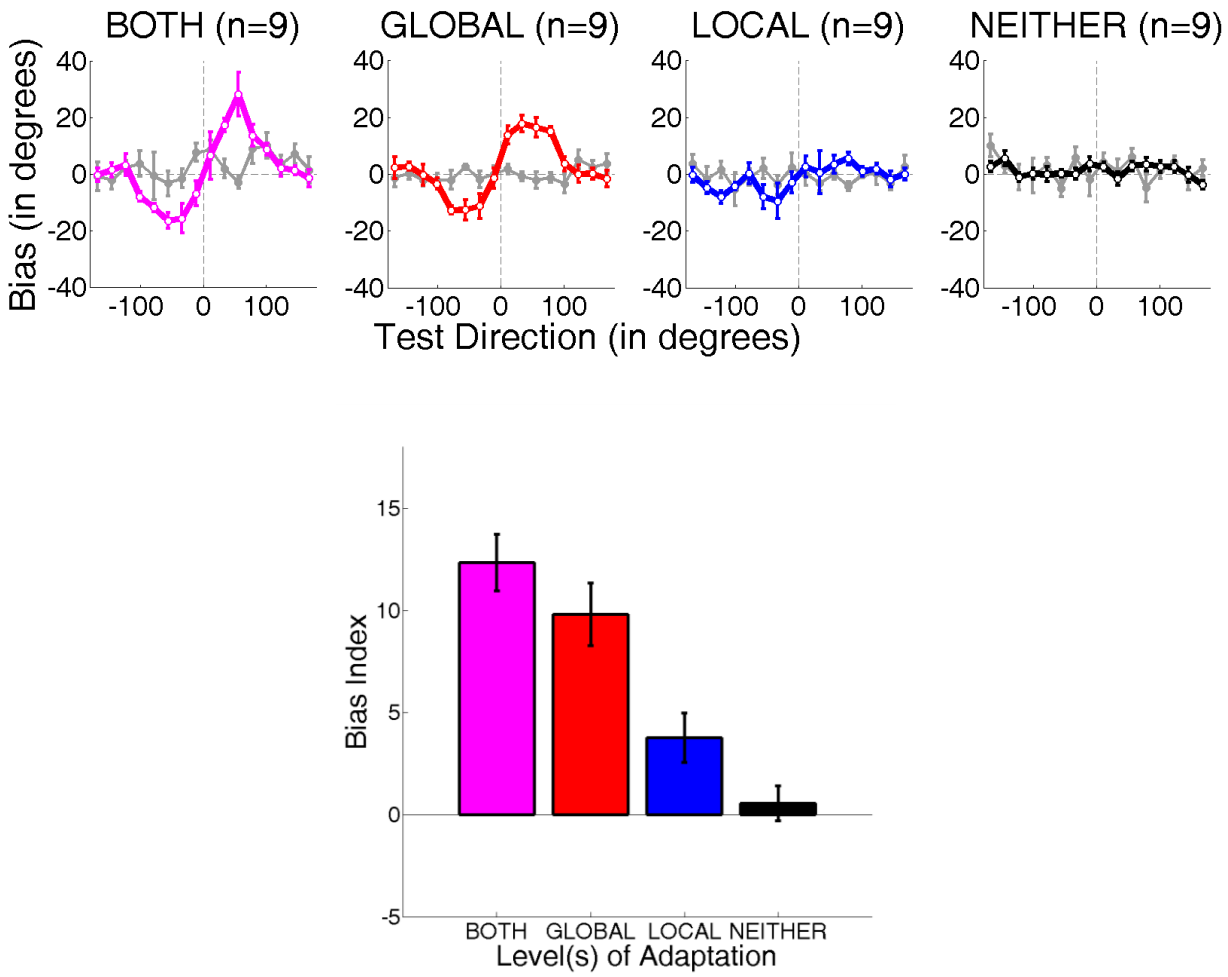

Figure 5.2. Results of Experiment 1A. Upper: Averaged perceptual biases of all four conditions in the Test-only session (thin, gray lines with solid symbols) and the Adapt-Test session (thick, colored lines with open symbols). Errorbars are ±1 between-subjects S.E.M. Lower: Averaged bias indices for the four conditions in the Adapt-Test session. Errorbars are ±1 between-subjects S.E.M.

The lower panel of Figure 5.2 shows the average bias index in the Adapt-Test session for the four conditions. For the Test-only session, average bias indices for Both, Global, Local, and

Neither were, respectively, 1.98, –.31, –.04, and .45. The relatively high average in the Both condition 1.98 was caused by one outlier (13.34; the second largest was 4.54). If we excluded this participant, the Both condition became .56, which was around zero. The overall trend is that observers did not show great bias in the Test-only session, which had been expected because there was no adaptation before they did the direction judgment.

We analyzed these data in the Adapt-Test session using a two-way ANOVA with two factors, test location (adapted vs. non-adapated) and global adaptation directions (one vs. five). Note that the two factors manipulated in this study corresponded to the presence and absence of adaptation at two levels, i.e., Global Adaptation and Local Adaptation . There was no significant two-way interaction ($F(1,32) = .164$, $p = .689$), and the main effects of both factors were found to be significant (Global Adaptation: $F(1,32) = 110.724$, $p < 10^{-11}$; Local Adaptation: $F(1,32) = 11.472$, $p = .002$). This suggests that adaptation at both levels contributed to the generation of DAE. In particular, the magnitude of perceptual bias was found to be much stronger in the Global condition (M = 9.83, SD = 2.42) than in the Local condition (M = 3.78, SD = 3.06; $F(1,32) = 25.458$, $p < 10^{-4}$), which was significantly stronger than null by itself (One-sample *t*-test against zero: $t(8) = 3.704$, $p = .006$). Perceptual bias in the Both condition (M = 12.36, SD = 2.77) was significantly stronger than that in the Global condition ($F(1,32) = 28.742$, $p = .043$). Finally, as expected, the bias index in the Neither condition was close to zero (M = .57, SD = 1.71). Results in Experiment 1A suggest that DAE depends on the processing level at which adaptation occurs. In particular, DAE was found to be much stronger when adaptation was introduced only at the global level than when the adaptation was introduced only at the local level.

Although the design of Experiment 1A allowed us to assess how level-specific adaptation may affect subsequent direction perception across the full 360° range of directions, there were several shortcomings in this experiment. First, there was a large variability in observers' responses in the direction judgment task. In the Test-only session where there should be no adaptation-induced bias, the variability of reported directions was large. Average SD of responses across all observers in the Test-only session was 21.7° (Range: 13.7° to 40.8°). Second, the action of turning the simulated dial was time-consuming, e.g., some observers took more than 4 seconds to respond in some trials. In these trials, effects of adaptation could be substantially weakened probably due to the fade of visual memory. Last, the four conditions were administered across subjects, what was not as ideal as using the within-subject design to compare the effect of level-specific adaptation within the same individual.

In order to obtain a more precise and reliable measure of the strength of DAE, we conducted Experiment 1B, which employed the same 2-by-2 design but a different psychophysical paradigm and a different task to address the above shortcomings. The goal of Experiment 1B was to narrow down the measurement of the strength of DAE at about $\pm 45^{\circ}$ away from the adapting direction. We used these directions based on the results of Experiment 1A in which the strongest DAEs were revealed with these directions. This approach has been used in previous psychophysical studies in quantitatively measuring the strength of DAE (Curran, et al., 2006; Wenderoth & Wiese, 2008).

Experiment 1B: Quantitatively measuring the strength of DAE caused by level-specific adaptation

Methods

Experiment 1B aimed to measure the point of subjective equality (PSE) of the upward direction after adapting to the oblique 45° (up-left or up-right) direction. In separate groups of subjects, we used the constant-stimuli method and the Psi adaptive-staircase (adaptive Psi) method (Kontsevich & Tyler, 1999) to measure the PSE. The adaptive-staircase method was implemented using the Palamedes Toolbox (Prins & Kingdom, 2009).

General stimulus parameters were identical to those used in Experiment 1A, unless otherwise specified. Five observers participated in the constant-stimuli experiment for monetary reward. Twelve observers participated in the adaptive-staircase experiment for course credits. In both experiments, the reference adapting direction r was either counterclockwise-45° or clockwise-45° relative to the upward direction. In the constant-stimuli experiment, nine test directions were chosen based on pilot experiments for each condition. They spanned a range of 60° and were evenly spaced near the upward direction. In the adaptive-staircase experiment, test directions were determined adaptively using the adaptive Psi method (Kontsevich & Tyler, 1999). Observers performed the direction discrimination task during the response phase, in which they were instructed to determine whether the overall motion direction of the test stimulus was to the left (counterclockwise) or to the right (clockwise) of the upward direction, which corresponded to the left and the right arrow keys, respectively.

In the constant-stimuli experiment, adaptation durations were the same as those used in Experiment 1A (Initial: 60s; Top-up: 10s). Observers completed the four conditions of experiments in separate sessions, with one condition assigned to each session. Order of

assignment was counterbalanced across observers. In each condition, observers made 144 direction discrimination responses (9 test directions × 16 trials). These 144 trials were blocked into 8 blocks of 18 trials. Two repetitions of the 9 test directions were randomly interleaved within a block. The counterclockwise and clockwise 45° adapting directions of r was alternated across the blocks, and there were 8 trials for each adapting direction in each condition. There was a 60-second rest period between every two blocks of trials. Observers completed two short practice blocks at the beginning of the first session for familiarization of the task.

In the adaptive-staircase experiment, initial and top-up adaptation durations were 45 and 6 seconds, respectively. The response phase was limited to 2 seconds only. Each participant completed 12 blocks of 16 trials, resulting in a total of 196 trials (48 trials for each of the 4 conditions). The number of global adapting directions (1 or 5) was alternated across blocks, while the reference adapting direction (counterclockwise-45° or clockwise-45°) was randomized across blocks, with the constraint that no three consecutive blocks had the same reference adapting direction. Within each block, trials with different test locations (adapted or non-adapted) were randomly interleaved. Specifically, trials of the Both and the Global conditions were interleaved within the same block, and trials of the Local and Neither conditions were interleaved within the same block. The PSEs of the upward direction for the 4 conditions were estimated on independent tracks of adaptive staircases. There was a 45-second rest period between every two blocks of trials. Observers completed two short blocks of practice before running the experimental blocks to familiarize themselves with the task.

Results

For the clockwise-45° adapting direction, the PSE of the upward direction was found to shift towards the clockwise direction, and vice versa for the counterclockwise-45°. As there was no

systematic difference between the two adapting directions, trials of these two adapting directions were combined in our analyses, so that responses would be labeled as "Adapting (+ve)" or "Opposite (-ve)".

For both the constant-stimuli and adaptive-staircase experiments, we used the cumulative normal distribution function to fit the psychometric curve for each observer. The two parameters, mean μ and standard deviation σ, were estimated based on each observer's responses in each condition. In particular, the μ parameter was taken as the estimate of the PSE, because it represents the test stimulus direction at which the observer's response was at chance (i.e., 50%) in judging with the test direction was to the left or the right of the upward direction.

In the constant-stimuli experiment, we fitted each observer's responses by least-square fitting. The left panel in Figure 5.3 shows the fitted psychometric curves of one observer (RH) for all four conditions. The R^2 values ranged from .955 to .9994, indicating that fitting was good for all observers in all conditions. The estimated parameter μ was taken to be the PSE estimate, which represents the strength of DAE at a test direction at about 45° from the adapting stimulus. The right panel in Figure 5.3 shows the PSE estimates for all observers in all conditions (averages as bars). The overall trend was similar to that found in Experiment 1A. In particular, DAE was stronger in the Global condition (M = 13.193°, SD = 2.33°) that in the Local condition (M = 9.81°, SD = 3.87°). Among the 5 participants, results from 4 participants show this direction of effect. DAE was the strongest in the Both condition (M = 19.78°, SD = 5.26°), and close to zero in the Neither condition (M = –1.06°, SD = 3.95°).

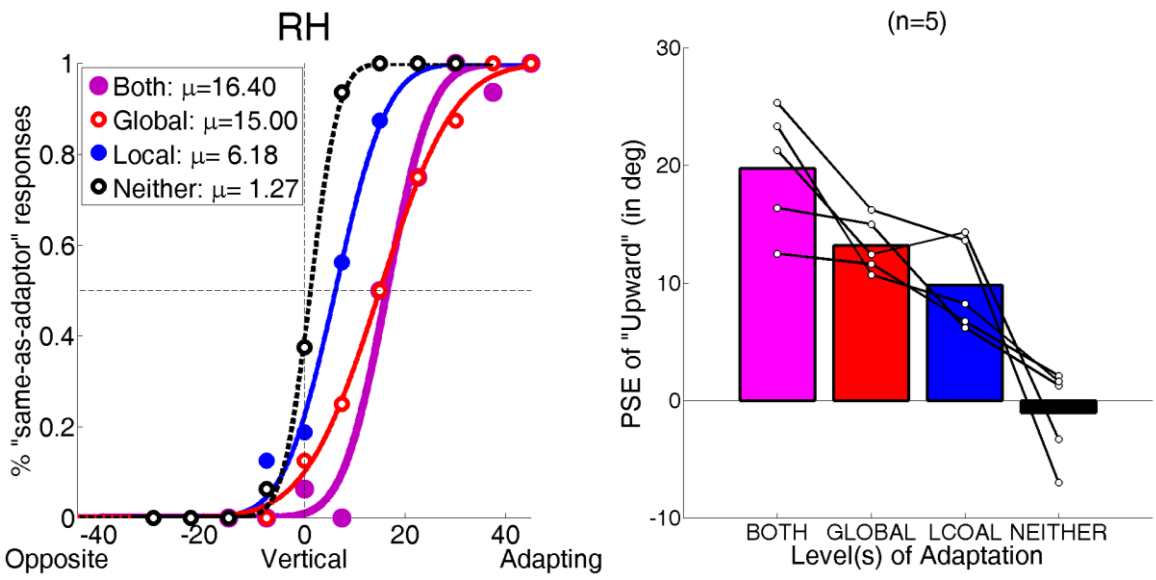

Figure 5.3. Results from the constant-stimuli experiment (n=5). Left: Psychometric curves of one observer (RH). For each observer, responses were fitted independent for each of the four conditions (Both: magenta, solid symbols thick line; Global: red, open symbols; Local: blue, solid symbols; Neither: black, open symbols dashed line). Right: PSE estimates for all observers in all conditions. Each open symbol indicates one PSE estimate in one condition. Individual observers' PSE estimates were connected by thin black lines. Each bar shows the PSE estimate for each condition, averaged across observers.

In the adaptive-staircase experiment, to be consistent with the fitting in the constant-stimuli experiment, we chose the cumulative normal distribution as the underlying psychometric function for the adaptive algorithm of the Psi method. The 50%-threshold was estimated based on this function as observers proceeded through the trials. For each observer, PSE was estimated using the Psi method for each condition. The left panel in Figure 5.4 shows the four tracks of adaptive-staircases, one for each condition, for one observer (TM). Estimates of PSEs (thick lines) typically stabilized early on. The fluctuation in the test directions (thin lines, open symbols)

was for better estimates of the slope of the psychometric curve, representing sensitivity, which was not the main concern for this experiment. The right panel in Figure 5.4 shows the PSEs, averaged across observers for the four conditions. Similar as the trend observed in the constant-stimuli experiment, the strongest DAE was observed in the Both condition (M = 13.43, SD = 7.68), followed by the Global condition (M = 11.02, SD = 6.30). DAE was found to be much weaker in the Local condition (M = 4.44, SD = 6.81), and was close to null in the Neither condition (M = 2.08, SD = 6.26).

We analyzed the data using a repeated-measures ANOVA, with the same two factors used in Experiment 1A (Global Adaptation and Local Adaptation). We did not find a significant two-way interaction ($F(1, 11) = .001$, $p = .971$) or a significant main effect of Local Adaptation ($F(1, 11) = 1.428$, $p = .257$). However, the main effect of Global Adaptation was significant ($F(1, 11) = 21.085$, $p = .001$, partial $\eta^2 = .657$). In particular, the Global condition revealed significantly stronger DAE than that the Local condition did ($F(1, 11) = 6.431$, $p = .028$, partial $\eta^2 = .369$). The magnitude of DAE in the Global condition was, similar to the effect in the Both condition ($F(1, 11) = 1.521$, $p = .243$, partial $\eta^2 = .121$). When compared with null, DAE in the Local condition was marginally significant ($F(1, 11) = 5.088$, $p = .045$, partial $\eta^2 = .316$).

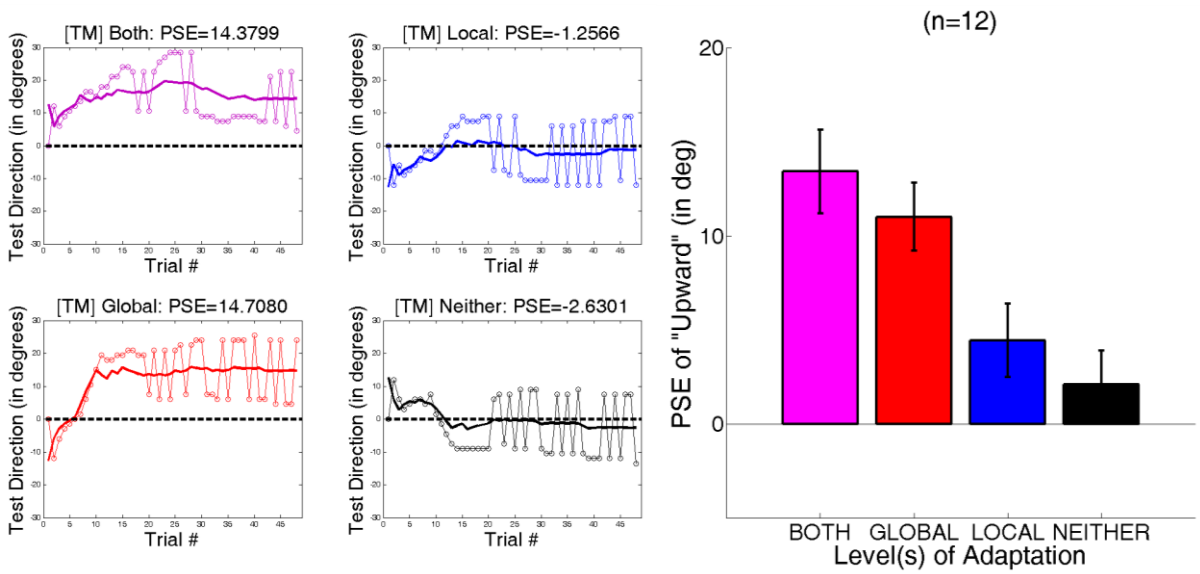

Figure 5.4. Results from the adaptive-staircase experiment (n=12). Left: Independent tracks of PSE estimates for the four conditions (Both: top-left magenta; Global: bottom-left red; Local: top-right blue; Neither: bottom-right black) for one observer TM. The thick line in each small panel traces through the estimated PSE across trials. Test stimulus direction presented at each trial is marked by open symbols, and connected by thin lines. Right: PSE estimates averaged across observers in all conditions. Errorbars are ±1 S.E.M.

Taken together, results from Experiment 1A and 1B provide evidence supporting that DAE is stronger when adaptation is specifically induced at the global level of motion processing than when induced at the local level. This difference can be explained by the adaptation of the global direction-selective units. It has been shown that the specific adaptation-induced changes in direction tuning found in neurons in MT (Kohn & Movshon, 2004) can lead to stronger perceptual repulsion (see chapter 6 for detailed explanation). The small DAE found in the Local condition was possibly caused by propagation of location-specific effects of adaptation (Xu, et al., 2008). While adapting to grating stimulus produces both repulsive aftereffects on both

motion direction (Schrater & Simoncelli, 1998) and speed (A. A. Stocker & Simoncelli, 2009), it is unclear which of these effects or both of them were propagated to produce the percept of DAE in the Local condition.

In the above two experiments, adapting orientation and test orientation remained the same at each location. The local motion signal at each location was confined to the single dimension that was perpendicular to the grating orientation (i.e., the "1D motion" termed in Amano, et al., 2009b). As a result, from adaptation to test, these local 1D motion vectors only changed their speeds, but not their directions. This local manipulation resembles that used by Stocker and Simoncelli (2009) to obtain significant repulsive speed aftereffect. Given that local test direction did not differ from the adapting one, the DAE observed in the Local condition may actually be driven by the propagation of multiple repulsive speed aftereffects across locations instead of repulsive direction aftereffects across locations.

Experiments 2 and 3 were conducted to pin down which local repulsive effect was propagated downstream to produce the overall DAE percept. In both Experiments, we attempted to reduce the effect of local speed repulsion and enhance the effect of local direction repulsion in the Local condition. If stronger DAE was observed, it would suggest that local direction repulsion contributes to the generation of a DAE percept. On the contrary, if DAE became weaker, it would suggest that local speed repulsion is necessary in generated a DAE percept.

Experiment 2: Manipulating test orientations

Methods

The stimulus and the procedure were exactly the same as the Local condition in Experiment 1B, except for the details described below. Experiment 2 used 5 global adapting directions that did not overlap. Elements presented at the test locations had been assigned a coherent global direction in the adaptation phase. Two new conditions, namely, the 45-degree condition and the Orthogonal condition, were included the Experiment 2. The 45-degree condition was designed so that, at each test location, the test speed was similar as the adapting speed. As explained in General Methods, at each test location j, the adapting speed u_j is given by

$$u_j = v \sin(r - \theta_j), \quad \ldots(1)$$

where r is the reference adapting direction, v is the global speed (same for both adaptation and test) and θ_j is the orientation of element j. The test speed u_j^t is given by

$$u_j^t = v \sin(t_i - \theta_j^t), \quad \ldots(2)$$

where t_i is the test direction for trial i, and θ_j^t is the orientation of element j during test. Across all conditions, the test direction t_i was always about 45° away from the adapting direction r (i.e., $t_i - r \approx \pm 45°$). In the Local condition in Experiment 1, adapting and test orientations were identical at each location (i.e., $\theta_j = \theta_j^t$). Therefore, the difference between the adapting r and the test directions t_i caused the adapting u_j and the test speeds u_j^t to be different. In the 45-degree condition of Experiment 2, we attempted to cancel out this speed difference by introducing a difference in orientation between adaptation θ_j and test θ_j^t. Each test element was turned 45° away from its original orientation during the adaptation phase, so that $\theta_j^t - \theta_j = \pm 45°$. The

direction of tilt (clockwise or counterclockwise) was determined based on the relative difference between the test direction t_i and the global adapting direction r for each trial. Formally,

$$\theta_j^t = \theta_j + sign(t_i - r) \times 45° \quad \ldots(3)$$

When t_i was counterclockwise to r, test orientations θ_j^t were tilted counterclockwise; when t_i was clockwise to r, test orientations θ_j^t were tilted clockwise. As a result,

$$\theta_j^t - \theta_j = 45° \approx t_i - r$$

$$\Rightarrow \quad r - \theta_j \approx t_i - \theta_j^t$$

$$\Rightarrow \quad u_j = v \sin(r - \theta_j) \approx v \sin(t_i - \theta_j^t) = u_j^t$$

Therefore, by tilting test orientations by 45°, we maintained similar speed at each location *j* between adaptation u_j and test u_j^t in the 45-degree condition. In the Orthogonal condition, each test element's orientation was turned 90° relative to its original adapting orientation. Because local motion detectors are known to be orientation-selective, probing the aftereffect with an orthogonal condition should result in minimal, if any, local aftereffects. This condition was designed to serve two purposes: 1) to assess how much the DAE found in the Local condition depends on orientation-specific adaptation; and 2) to provide a bench mark in comparison with the effect found in the 45-degree condition. We ran the constant-stimuli experiment on the same 5 observers in Experiment 1B for both the 45-degree and the Orthogonal conditions.

Simulation

If local speed repulsion was necessary in the generation of the DAE observed in the Local condition in Experiment 1, removing it using this orientation-tilting method should reduce the strength of the DAE. In order to test this hypothesis, we ran a toy simulation by applying the speed-repulsion model developed by Stocker and Simoncelli (2009) to each test location. Their

model consisted of two speed-repulsion components, $\delta_{j,ND}$ and $\delta_{j,D}$, each of which can be described as a derivative of Gaussian

$$\delta_{j,ND} = -w_{ND}\left(\frac{\left(0-u_j^t\right)}{s_{ND}^3\sqrt{2\pi}}\exp\left(\frac{-\left(u_j^t-0\right)^2}{2s_{ND}^2}\right)\right), \quad \ldots(4)$$

$$\delta_{j,D} = -w_D\left(\frac{\left(u_j^a-u_j^t\right)}{s_D^3\sqrt{2\pi}}\exp\left(\frac{-\left(u_j^t-u_j^a\right)^2}{2s_D^2}\right)\right), \quad \ldots(5)$$

$$\delta_j = \delta_{j,ND} + \delta_{j,D} \quad \ldots(6)$$

where δ_j is the total amount of repulsive bias in speed for element j, u_j^a is the adapting speed for element j, u_j^t is the test speed for element j, and w and s are the two parameters that control, respectively, the amplitude and the range of effect of speed adaptation. The two components, δ_{ND} and δ_D, refer to, respectively, the non-directional and the directional repulsive effect of speed adaptation. The non-direction component $\delta_{j,ND}$ centers at 0 and does not depend on adapting speed, whereas the directional component $\delta_{j,D}$ centers at the adapting speed u_j^a

We fixed the values for all four parameters ($w_D = 1.08, w_{ND} = -0.36, s_D = 0.8$ and $s_{ND} = 1.2$) in our simulation. The range parameters s_D and s_{ND} were chosen to match those found by Stocker and Simoncelli in their empirical fitting. The amplitude parameters w_D an and w_{ND} were chosen to fit the results of Experiments 1B for the average of the five observers. We roughly maintained the same ratio between w_D and w_{ND} (i.e., 3: –1) as found in Stocker and Simoncelli's experiment.

To simulate an overall DAE percept based on local speed repulsion, we generated $N = 100$ local-motion elements, each with a randomly-assigned orientation θ_j. We then set the global

adapting direction r as 45°, and the global test direction t as 90°, and applied equations (1) and (2) to obtain the adapting $u_j$and test speeds u_j^t (left and middle columns of Figure 5.5). With these speed values, we then applied equations (4) and (5), and then (6) to compute the local speed repulsion for each location.

We computed the perceived speed $\hat{u}_j$ at each location j by adding the repulsive speed bias to the veridical test speed

$$\hat{u}_j = u_j^t + \delta_j \quad \ldots (7)$$

The local motion direction at each location is simply perpendicular to the orientation of the Gabor element at that particular location. The positive or negative sign of the perceived direction was determined by the sign of $\hat{u}_j$. Note that, in this toy simulation, we did not apply any direction repulsion at each location. As a result, we obtained a perceived speed and direction at each location, which would give us a set of perceived local motion vectors $\{\widehat{\boldsymbol{u}}_j\}$ (right columns of Figure 5.5).

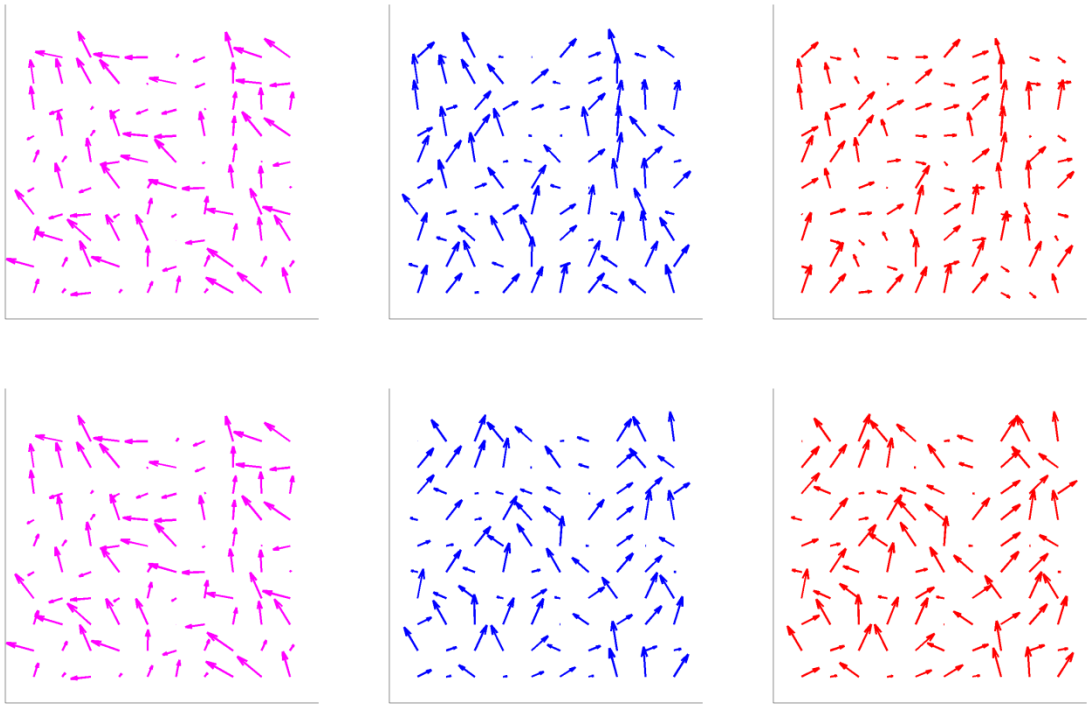

Figure 5.5. Illustration of motion flows in simulation. (Left column) Adapting velocities, which are identical across conditions. (Middle column) Test velocities. Local directions depend on element orientations, which are not shown in this illustration. (Right column) Perceived velocities computed purely based on propagation of local speed repulsion. (Top row) The Local condition in Experiment 1. Local orientations remain unchanged between adaptation and test, and, thus, local directions remain unchanged. (Bottom row) The 45-degree condition in Experiment 2. The changes in local adapting and test directions are determined purely by changes in orientations. No direction repulsion was applied in the simulation.

We then obtained the overall perceived direction by computing the angle of the averaged vector of $\widehat{\widehat{\boldsymbol{u}}}_j$ across all locations. Finally, the overall DAE was taken to be the difference between the test direction (i.e., 90°) and the overall perceived direction.

$$DAE = \arctan\left(\frac{1}{N}\sum_j \widehat{\widehat{\boldsymbol{u}}}_j\right) - t$$

To simulate the original Local condition in Experiment 1, we set the test orientation to be the same as the adapting orientation at each location, i.e., $\theta_j^t = \theta_j$. As the amount of speed bias δ_j depends on the amplitude parameters w_D and w_{ND}, we fixed them so that the overall DAE matched with the average of the five observers (~10°). To simulate the 45-degree condition in Experiment 2, we set the test orientation to be 45° away from the adapting orientation at each location based on equation (3). To simulate the Orthogonal condition in Experiment 2, we set the test orientation to be orthogonal to (90° away from) the adapting orientation.

Results

Simulation results are shown in Figure 5.6 (right panel). Over 5000 runs, averaged DAEs for the Local, the 45-degree and the Orthogonal conditions, were found to be 9.29°, 2.06° and 0.60° respectively, all repulsive relative to the adapting direction. Purely based on propagation of local speed repulsion, our simulation predicts that DAE is greatly reduced if test orientations are tilted from 45° relative to their corresponding adapting orientations.

For the results on human observers, we performed the same analyses as we did in Experiment 1B. Results are shown in Figure 5.6 (left panel), together with the results from the Local condition obtained in Experiment 1B (labeled as "Same" to denote the same orientations between adaptation and test) for easy comparison. The shift of PSE was significantly modulated by test orientation (repeated-measures ANOVA: $F(2, 8) = 20.978$, $p = .001$). In particular, DAE in the

45-degree condition (M = 3.44, SD = 2.84) was significantly weaker than that in the Same condition ($F(1, 4) = 15.611$, $p = .017$). This trend was consistently observed across all five observers. DAE was weak in the Orthogonal condition (M = 1.81, SD = 1.16). Across all five observers, DAE was slightly stronger in the 45-degree condition than in the Orthogonal condition, although, with only five observers, the current analysis lacks statistical power to test the significance of such different. Overall, results from human observers can be predicted by a simple model that produces DAE purely based on local speed repulsion.

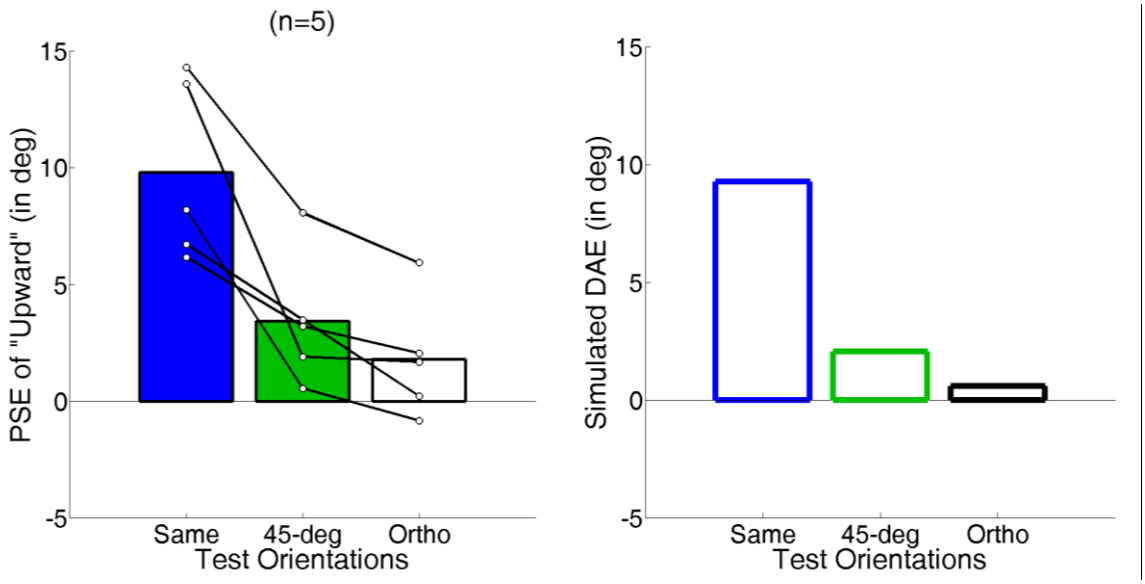

Figure 5.6. Results of Experiment 2 and simulation. (Left) Results of the Local condition (labeled as "Same") in Experiments 1A and 1B were replotted in the figure for easy comparison. Results from the constant-stimuli experiment for the three conditions. Each open symbol indicates one PSE estimate in one condition. Individual observers' PSE estimates were connected by thin black lines. Each bar shows the PSE estimate for each condition, averaged across observers. (Right) Results from the simulation of local speed repulsion. Hollow bars show the simulated DAE for each condition averaged over 5000 runs of simulation.

In summary, results from Experiment 2 suggest that the percept of DAE in the Same condition may be mainly driven by the propagation of local speed repulsion. This suggests that the percept of DAE may not be due to integrating repulsively-biased local motion directions. However, DAE strength does seem to depend on local test orientations. One possible explanation is that what is integrated is the distribution of activities in low-level motion detectors, which depends on stimulus orientation, but does not generate local percepts. This is possible because local motion signals carried by drifting Gabor elements are inherently ambiguous in representing the global motion direction (Marr & Ullman, 1981). Therefore, this multiple-aperture stimulus may prevent the visual system from reading out local motion percepts from the activities of low-level motion detectors at each location, as that would interfere with the global percept of motion. If so, disambiguating the local motion signals may facilitate the read out of local illusory signals, which could then be integrated to produce a global percept of DAE. Experiment 3 was designed to test this hypothesis by using the plaid stimulus at each location, which contains an unambiguous local motion vector that is identical to the global motion vector. Again, we focused on the original Local condition. Because we maintained the same global speed and only altered global direction between adaptation and test, using plaid at each location should lead to little local speed repulsion and allow more room for local direction repulsion (if any). If adapting to unambiguous local signals could produce illusory local repulsive signals that can be later integrated to produce stronger percept, we should be able to observe stronger DAE in the Local-Plaid condition than in the Local-Gabor condition. However, if we observe similar or even weaker DAE for the plaid stimulus, it would suggest that integration of local illusory repulsive signals does not contribute to the percept of DAE.

Experiment 3: Disambiguating local motion signals to produce local percepts

Methods

The stimulus and the procedure were exactly the same as those in previous Experiments, except for the details described below. In Experiment 3, we created the Plaid condition based on the original Local condition in Experiment 1. At every location, we superimposed two identical Gabor elements with a 90° difference in orientation to create a plaid. Overall contrast of each plaid element was maintained at the same level as the previous two Experiments by halving the contrast of each of the two underlying Gabors. We ran two separate experiments for the Plaid condition to obtain quantitative measures of the the strength of DAE, using the constant-stimuli method and the adaptive-staircase method, respectively. The same five observers as the previous Experiments participated in the constant-stimuli experiment, and 18 observers participated in the adaptive-staircase experiment. In the adaptive-staircase experiment, each observer ran both the Gabor and the Plaid conditions in the same session.

Results

We analyzed the data in the same way as in previous Experiments. Results from the Plaid condition were compared with perform using Gabor elements for the five participants using the constant-stimulus paradigm and for the 18 participants using the adaptive-staircase paradigm. As shown in Figure 5.7, DAE was weaker in the Plaid condition than that in the Gabor condition. For the measures from the constant-stimuli paradigm, DAE was significantly weaker in the Plaid condition (M = 1.57, SD = 3.59) than in the Gabor condition (M = 10.78, SD = 5.53; Paired-samples t-test: $t(4) = 5.835$, $p = .0043$). Importantly, this trend was consistently observed across all five observers. For the measures from the adaptive-staircase paradigm, the PSE shift in the Plaid condition (M = 6.40, SD = 2.83) was, again, smaller than that in the Gabor condition (M =

8.58, SD= 4.45). A paired-samples t-test reveal a marginally significant effect ($t(17) = 2.04$, $p = .058$), and the "Gabor > Plaid" trend was observed in 13 out of 17 observers (72%). This finding suggests that, in the absence of local speed repulsion, perceived DAE becomes much weaker, even when there is room for the generation of local direction repulsion.

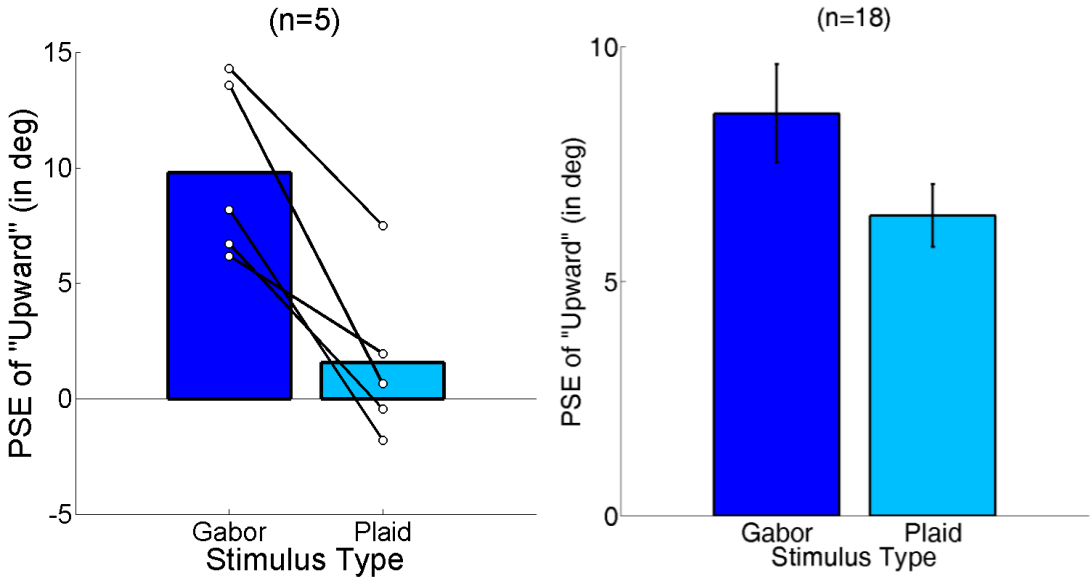

Figure 5.7. Results of Experiment 3. Bars represent averages across observers in each condition. Errorbars are ±1 between-subjects S.E.M. Left: PSE shift obtained from the constant-stimuli experiment ("Gabor" was replotted from results of Experiment 1B). Each open symbol indicates one PSE estimate in one condition. Individual observers' PSE estimates were connected by thin black lines. Right: Results from the adaptive-staircase experiment.

General Discussion

In a series of psychophysical experiments, we found that 1) DAE was stronger when only the global level adapted than when only the local level adapted; 2) DAE arising from prorogating adaptation-induced local effects was orientation-specific; and 3) DAE arising from local adaptation was weaker when local motion signals were unambiguous. These results suggest that DAE depends on the level(s) at which adaptation occurs. In particular, our findings suggest that the DAE that arises from local adaptation is the result of integrating illusory speed signals that are orientation-specific, but not local illusory directional signals.

Our results appear to be inconsistent with the conclusion drawn by Curran and colleagues (2006), that DAE is mainly driven by the adaptation of local-motion detectors. One possible factor is that they used a random-dot stimulus for both adaptation and test, which was different from the multiple-grating stimulus used in the present study. Although the stimulus difference likely contributes to the divergent results to a certain extent, we doubt that it is the main reason since Schrater and Simoncelli (1998) found that DAE did not depend on spatial pattern (dots or gratings) or even on spatial frequency. To explain the discrepancy of findings, we noticed the difference in defining "local" and "global" levels of motion processing in the two studies. In Curran et al.'s study, they first measured how DAE depended on motion speed, so that they have a lookup table from motion speed to DAE strength. Then, in arbitrating between the contribution of local and global adaptation in producing DAE, they proposed a "local" model and a "global" model for generating DAE. The "local" model first looks up DAE strength for each speed level present in the stimulus (which was taken as "local speed"), and then averages the DAE strengths. The "global" model first evaluates the "global" speed of the whole stimulus by averaging all "local" speed levels, and then looks up one DAE strength. In the present study, local and global

levels of motion processing were basically defined by receptive-field size (i.e., local: small, for each element; global: large, for the whole pattern) and the type of computation performed (i.e., local: motion-energy computation; global: spatial integration). We acknowledge that the definition employed by Curran and colleagues has its root in previous psychophysical studies on speed perception, and ours more in neurophysiological and computational studies. After all, as the rationales of the two studies are quite different, it would be hard to directly compare the apparently contradictory conclusions drawn in the two studies. It is possible that the motion system produces DAE based on many factors, and speed and spatial locations may just be tapping different sides of the same system. If one focuses on local speed representation, findings from both studies are not exactly opposite: both studies support that the percept of DAE depends on local speed representation. In their study, the integration of speed-dependent DAEs was found to better predict perceived DAE. In our study, we found that the propagation of local speed repulsion was necessary and sufficient in the generation of perceived DAE.

However, results of Experiment 1 are consistent with those reported in some studies (Jin, et al., 2005; Kohn & Movshon, 2004). Specifically, the trend that the DAE found in the Global condition was stronger than that found in the Local condition can be explained by the difference in adaptation-induced tuning changes found in neurons at V1 (Dragoi, et al., 2000; Muller, Metha, Krauskopf, & Lennie, 1999) and MT (Kohn & Movshon, 2004) (more details will be provided in the next chapter). After adaptation, the attractive shift of preference in low-level motion-sensitive neurons weakens the DAE caused by gain reduction, while the repulsive shift of preference in high-level motion neurons strengthens the DAE. Nevertheless, one may question whether it is appropriate to label different “repertoires” of adaptation-induced tuning changes by processing levels. In a recent study (Wissig & Kohn, 2012), it has been shown that neurons in V1

can also demonstrate the types of tuning changes, particularly attractive shift of preference, found in neurons in MT. However, the adapting stimulus that triggered the attractive shift of preference in V1 neurons was much larger than the receptive fields of those neurons. These adapting stimuli may have adapted neurons at MT, which could send feedback signals to V1 neurons and cause them to shift of their preferences towards the adaptor.

Taken together, findings from Experiments 2 and 3 demonstrate that manipulation of local features (i.e., orientation and local ambiguity) can lead to changes in the percept of the DAE that generated from local adaptation. In line with findings from our previous study (Lee & Lu, 2012) and others (Solomon, et al., 2004; A. A. Stocker & Simoncelli, 2009; Xu, et al., 2008), these results suggest that effects of adaptation at low level of processing can be propagated downstream to affect later stages of processing.

While converging evidence suggests that adaptation in the processing hierarchy cascades through the levels, it is unclear how such propagation is carried out. Considering visual processing in the cortex (i.e., starting from V1), an intuitive proposal is that low-level adaptation creates local illusory signals that are then integrated by later stages of processing to produce an aftereffect percept. Indeed, such "integration-of-illusions" account is appealing, because of its parsimony (because it builds on how the system integrates "real" signals) and its power in explaining aftereffect phenomena. For example, it can explain why the percept of motion aftereffect (MAE) is always integrated after adapting to bidirectional transparent motion (Vidnyanszky, et al., 2002), as well as some of our previous findings (Lee & Lu, 2012; and Chapter 4). However, findings from the present study may provide some counterevidence to this account. The manipulation in both Experiments 2 and 3 aimed to strengthen illusory local DAE signals, so that, if they could be integrated by some high-level units in later stages of processing, the final DAE percept could

be strengthened. However, the results went in the opposite direction: strength of DAE did not increase, and even appeared to decrease when local DAEs were strengthened. This seems to go against the proposal that low-level effects of adaptation are propagated via integration of local illusory signals. If it is not the local illusory signals that are integrated, how are effects of adaptation propagated downstream?

One possibility is that neural activities at early stages of processing do not contribute in producing conscious perception, so that they do not directly produce "local illusory signals". However, activity patterns at these early levels are later integrated internally to produce altered activity patterns in neurons at higher levels, from which a percept of, say, motion direction is read out. For example, local motion detectors are orientation-selective, and, therefore, so is the effect of adaptation at the local level. Probing the local level with test orientations that are far away from the adapting orientation may reduce the effect of adaptation on altering neural activity. The integration of these less-altered activity patterns are then fed to later stages of motion processing, and, thus, produces a less biased, weaker DAE percept. In short, our findings seem to suggest that there may be no perceptual readout at the local level of motion processing, at least in the case of the generation of DAE. This idea is consistent with the idea that conscious perception resides at higher level of processing than at low levels (Crick & Koch, 1995; Kanwisher, 2001). It is also consistent with a previous adaptation study in which adaptation to oriented gratings with imperceptibly high spatial frequency was found to produce the tilt aftereffect (He & MacLeod, 2001).

How can our findings be reconciled with the idea about integrating illusory signals to produce perceptual aftereffect? One possibly interpretation is that the existence of "local illusory signals" is just a shorthand representation of the population response pattern produced by adapted local

motion detectors during the test. In particular, for motion adaptation, when the test stimulus is stationary, population response at the local stage of processing can be represented as a "illusory" motion signal that is opposite to the adapting one. In our previous studies presented in Chapter 4, a model that integrates illusory signals generated at multiple locations can explain the perception of static MAE. The apparent discrepancy will be resolved if one views the propagation mechanism as propagating neural signals instead of "illusory motion signals".

Our findings do not rule out the possibility that that the integration mechanism for generating DAE is different from that for generating static MAE. As noted in many previous psychophysical studies (for a review, see Mather, et al., 2008), stationary and dynamic test stimuli may actually be probing different levels of motion processing. DAE, by definition, must be probed by a dynamic test stimulus. It is possible that the level at which aftereffect percepts are read out depends on the test stimulus: when the test is stationary, local illusory percepts are first read out and then integration; when the test is dynamic, an aftereffect percept is read out only at the global level from the integrated neural responses. However, this flexible propagation strategy may be less general and parsimonious than a "late-readout" strategy.

Chapter 6

A Multilayer Network Model for Hierarchical Motion Adaptation

Introduction

Visual adaptation produces remarkable perceptual effects. Previous three chapters provide psychophysical evidence to unveil the complex interaction between adaptation-induced effects (i.e., motion aftereffect and repulsion effect) and the level of motion processing at which adaptation occurs. However, it remains unclear what basic neural mechanisms underlie visual adaptation, and how these adaptation-induced neural changes at different levels of motion processing are related to different perceptual aftereffects. In this chapter, I aim to provide a computational account for hierarchical motion adaptation via preliminary simulations. I will first summarize physiological findings of the adaptation-induced changes on neural mechanisms, and then make a connection between adaptation-induced neural changes with observed perceptual phenomena.

Previous neurophysiological studies have found that, after prolonged stimulation, neurons at different levels of motion processing changed their response characteristics differently. In V1, the bell-shaped tuning curve of an orientation-selective neuron was found to undergo systematic adaptation-induced changes (Dragoi, et al., 2000; Muller, et al., 1999). Previous research has categorized these changes into three main types (Gilbert & Wiesel, 1990; Jin, et al., 2005; Schwartz, et al., 2007): 1) reduction of gain or responsivity, 2) increase in or broadening of tuning width, and 3) repulsive shift of preference relative to the adapting stimulus.

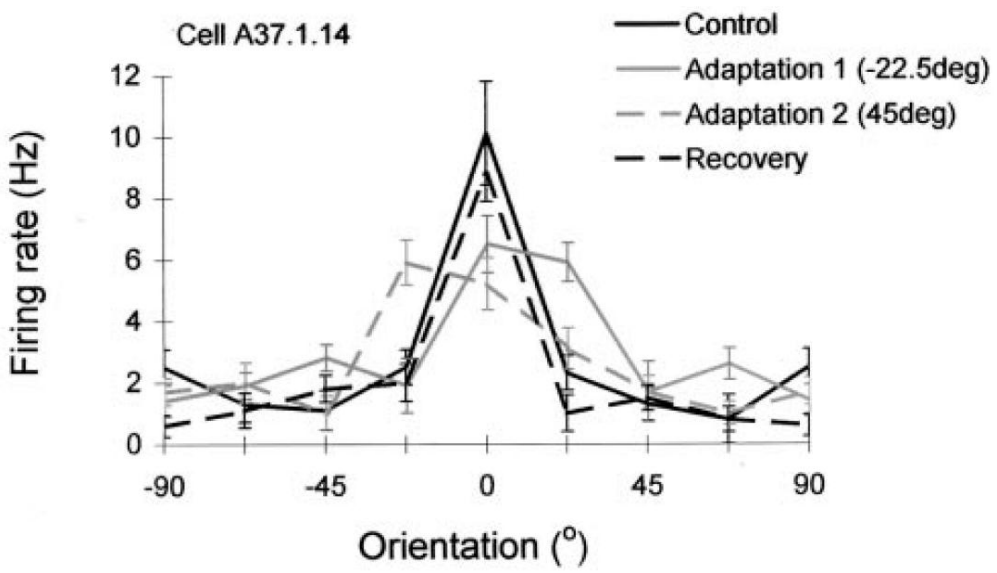

Figure 6.1. Adaptation-induced tuning changes in V1 (adopted from Dragoi et al., 2000). Gray lines indicate post-adaptation tuning. This neuron original preferred orientation is normalized to be 0°. When adapting orientation was –22.5° (solid gray), tuning peak shifted to the positive side, which is a repulsive shift away from the adaptor. When adapting orientation was 45° (dashed gray), tuning peak, again, shifted away from the adaptor towards the negative side. In general, tuning was also broadened and responses were suppressed.

Figure 6.1 showed an example from the study by Dragoi and colleages (2000) to illustrate these characteristic changes. It is worth noting that not all neurons demonstrate the same degree of these changes. A qualitative pattern is that, the closer the neuron's preferred stimulus is to the adapting stimulus, the stronger the adaptation-induced changes are observed. Nonetheless, motion sensitive neurons do not always demonstrate the above-mention types of adaptation-induced changes. At higher-level of motion processing (e.g., MT), adaptation-induced changes

were not exactly same as the changes in V1 (Kohn & Movshon, 2004): after prolonged viewing of motion stimulus, neurons in MT of the primate brain were found to demonstrate 1) gain reduction, which was similar to most V1 neurons showed, 2) narrowing of tuning width, and 3) attractive shift of preference towards the adaptor. Figure 6.2 showed the results from Kohn and Movshon's study to illustrate the adaptation-induced changes at MT level.

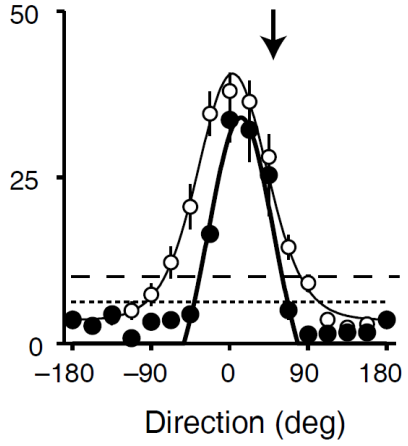

Figure 6.2. Adaptation-induced tuning changes in MT (adopted from Kohn & Movshon, 2004). The arrow indicates the adapting motion direction. Thin line (open symbols) indicates direction tuning before adaptation. Thick line (closed symbols) indicates direction tuning after adaptation. As shown, after adaptation, tuning peak shifted towards the adapting direction, tuning width was narrower, and response was moderately reduced.

To relate these adaptation-induced neural changes to perceptual aftereffect, I will use repulsive aftereffects as an example to show the qualitative connection. Repulsive aftereffects are believed to primarily arise from gain reduction at the adapting stimulus (Jin, et al., 2005; Schwartz, et al., 2007). An intuitive explanation is that neural responses are redistributed after adaptation (Mather

& Moulden, 1980): neural units that prefer the adapting stimulus are suppressed, while other units preferring other stimulus values were relatively enhanced, resulting in a population response pattern that is peaked at (and, thus, biased towards) stimulus values that are further away from the adaptor. Indeed, several analyses have demonstrated that gain reduction is the main cause of perceptual aftereffects, particularly in driving perceptual repulsion, and can be computationally described as a process of divisive normalization (Carandini & Heeger, 2012; Heeger, 1992). Since both V1 and MT neurons reveal gain reduction induced by adaptation, perceptual repulsion effect would be observed regardless whether an experimenter used simple motion (i.e., one drifting grating element) to probe low-level motion processor or global motion stimulus (i.e., multiple-aperture stimulus used in this thesis) to probe high-level processing. However, the magnitude of the repulsion effect may vary due to the different neural changes at different processing levels, i.e., shift in tuning peaks relative to the adaptor. The adaptation-induced repulsive shift of tuning peaks in V1 neurons would result in perceptual "attraction" if they were manifested independently, as they would bias the population response pattern *towards* the adaptor. This would weaken the perceptual repulsion caused by gain reduction in neurons when all three types of changes were in action (Gilbert & Wiesel, 1990; Jin, et al., 2005; Schwartz, et al., 2007). On the contrary, Kohn and Movshon (2004) showed that the attractive shift of tuning peaks in MT neurons would bias the population response pattern *away* from the adaptor. Assuming the perceptual effects caused by other types of changes remain unchanged, the attractive shift of tuning peaks found in neurons at this higher level of processing should lead to a stronger PR effect. This can explain the pattern of results obtained in the previous chapter, which demonstrate that global-level adaptation leads to stronger direction aftereffect (DAE) than local-level adaptation does.

Although one can relate adaptation-induced neural changes to perceptual aftereffect, a fundamental question still remains unanswered: *Why* do different processing levels (e.g., V1 and MT) undergo different adaptation-induced neural changes? Previous computational models have focused on the functional goals of adaptation, such as viewing adaptation as a form of efficient coding (Barlow, 1990), rescaling and recalibration (C. W. Clifford, et al., 2000), and enhancement of signal-to-noise ratio around the adapting stimulus (A. Stocker & Simoncelli, 2006). However, none of these theories address the issue of hierarchical adaptation, particularly why neural units at different processing levels appear to adapt differently. One appealing approach is predictive coding (Rao & Ballard, 1999), which states that, in the sensory system, lower-level units propagate prediction errors to higher levels, and feedback from higher levels contains information about the system's current prediction about the sensory world. This approach has recently been demonstrated to be powerful in predicting contextual effects, as well as low-level adaptation phenomena (Lochmann, Ernst, & Deneve, 2012). However, no models so far can simultaneously demonstrate the two types of adaptation-induced tuning changes found in V1 and MT, respectively.

Recently, Stevenson and colleagues (2010) have proposed an interesting computational account for neural adaptation. Given that neural activity is a combination of fluctuations in neural excitabilities (noise) and activation caused by stimulus in the world (signal), the authors suggested that adaptation-induced neuronal changes result from the process of the postsynaptic neuron actively estimating the presynaptic neuron's excitability based on their input signals. To minimize the influence of internal noise due to changing presynaptic neural properties such as excitability, the postsynaptic neuron aims to recover the sensory-driven signal via normalizing the input by the estimated excitability of the presynaptic neuron. Based on this theory, the

authors illustrated that a simple two-layer network can predict a range of adaptation-induced neural changes, including the typical V1 tuning changes found in neurophysiological studies (Dragoi, et al., 2000).

One of the key assumptions in their model is that sensory drive is sparse (i.e., presence of stimulus particularly preferred by a neuron is relatively rare). One implicit assumption in their model is that sensory drive is temporally independent. Although this sparseness assumption may match with the statistics of sensory-driven activities in V1 neurons, it does not match to statistics in natural dynamic scenes. For example, Roth & Black (2007) showed that optical flows in natural motion scenes contain a great deal of temporal regularity by measuring statistics in natural environment. In particular, motion flows tend to be temporally smooth (i.e., distribution of temporal derivatives peaks at zero). It is possible to incorporate this piece of information about global motion statistics into the Stevenson et al.'s general framework of adaptation, so we can explain the differences in adaptation-induced neural changes found in neurons in V1 and MT. This chapter aims to formulate a model for hierarchical motion adaptation by extending Stevenson et al.'s framework, in order to simultaneously predict adaptation-induced tuning changes in both V1 and MT. We added one MT layer after the V1 layer in their original model, and incorporated the temporal smoothness assumption in units at the MT layer. Figure 6.3 showed the basic architecture of the hierarchical model. Preliminary simulations revealed that the model predicts the adaptation-induced changes qualitatively similar to those reported in previous neurophysiological studies. This suggests that the level-specific adaptation-induced changes may be explained as neurons at different processing levels employing different assumptions about sensory statistics.

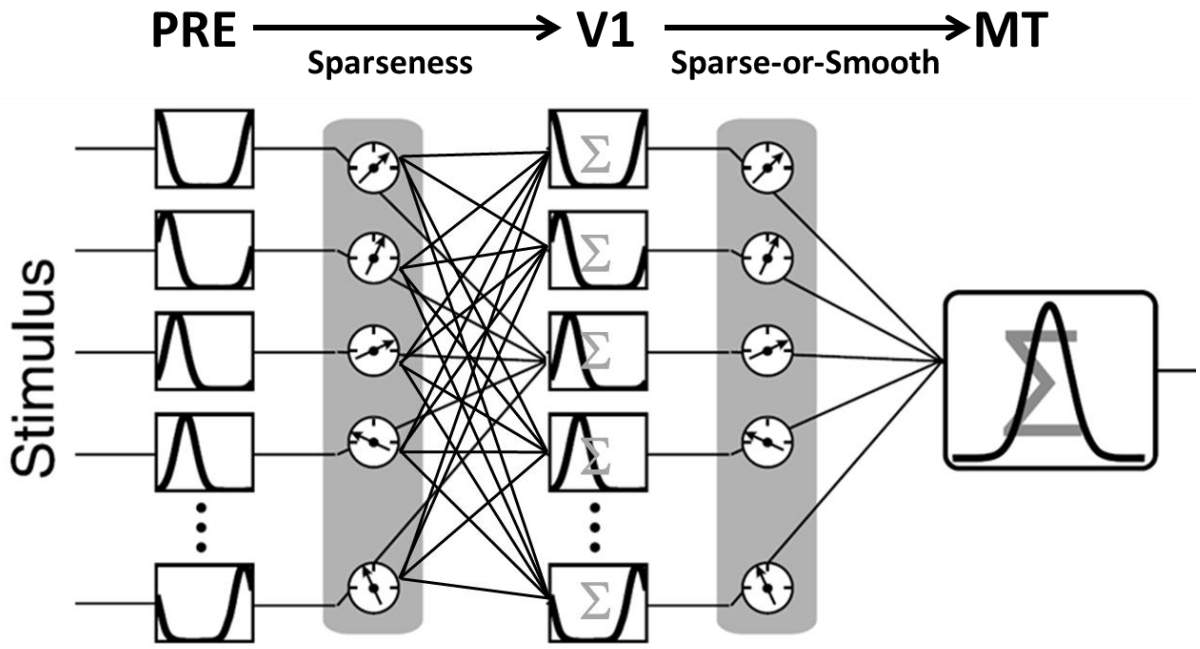

Figure 6.3. Schematic architecture of our three-layer feedforward network model (adapted from Stevenson et al.'s (2010)). Boxes represent neural units. The PRE units are directly tuned to stimulus direction, which determines the sensory drive. The gain symbol connected to each PRE unit represents fluctuation of total excitability for that unit. Therefore, the output from each PRE unit is a product between stimulus-based sensory drive and internal noise of excitability. The gray region (between PRE and V1) represents the computational problem faced by V1 units: they estimate the excitability of each PRE unit based on the presynaptic activity. The word "Sparseness" on top of that gray region refer to the statistical assumption V1 units hold in order to solve this estimation problem. Output signals from PRE are then normalized by their corresponding estimated excitabilities, and combined by each V1 neuron based on their tuned weighted sum. From V1 to MT, the structure is similar, except that we assume MT units employ the "Sparse-or-Smooth" assumption in their excitability estimation task.

Model Formulation

The model is inspired by the two-layer model proposed by Stevenson et al.'s (2010). Our model is a three-layer, feedforward network, as shown in Figure 6.3. We assume the connection between units across any two layers to be a simulated synapse. In general, the computation at each synapse is identical: at each time point, the postsynaptic unit estimates the excitability of the presynaptic unit based on the presynaptic input, and produces a postsynaptic output by normalizing the presynaptic input with the estimated excitability. The difference across levels lies in the different statistical assumptions about sensory drive held by units at different levels, based on which they estimate the excitability of the presynaptic units.

We refer to the three layers in the model as PRE (PREcortical), $V1$ and MT in reference to different stages of motion processing. This section is organized as follows. First, we describe how units at the PRE level produce non-adaptive output activities based on stimulus motion directions. Second, we describe the connection between the PRE and the $V1$ layers, and how $V1$ units produce adaptive outputs via excitability estimation process. Third, we describe the sequential learning algorithm by which $V1$ units adaptively estimate the excitabilities of PRE units based on their activities. Last, we describe how normalized outputs from the $V1$ units are fed to the MT units as input, and how MT units estimate the excitabilities of $V1$ units based on different statistical assumptions on the sensory environment.

The PRE level

Similar to Stevenson et al.'s model, the output (S^t) of each PRE unit at time t is defined as the product of the unit's total excitability (G^t) and the sensory drive (D^t) at that time:

$$S^t = G^t D^t \quad \ldots (1)$$

The sensory drive D^t can be understood as the "signals" that represent the stimulus in the sensory world. It is high when the unit's preferred stimulus is present and low when absent. G^t can be understood as the noise in the neural system that "contaminates" the signals. Thus, S^t represents the noisy neural activities observed at the first, PRE layer. In essence, the computational goal of each unit is to estimate the noisy excitability G^t, so that its influence on the presynaptic activity can be discounted and the postsynaptic unit can focus on responding to the true sensory drive D^t.

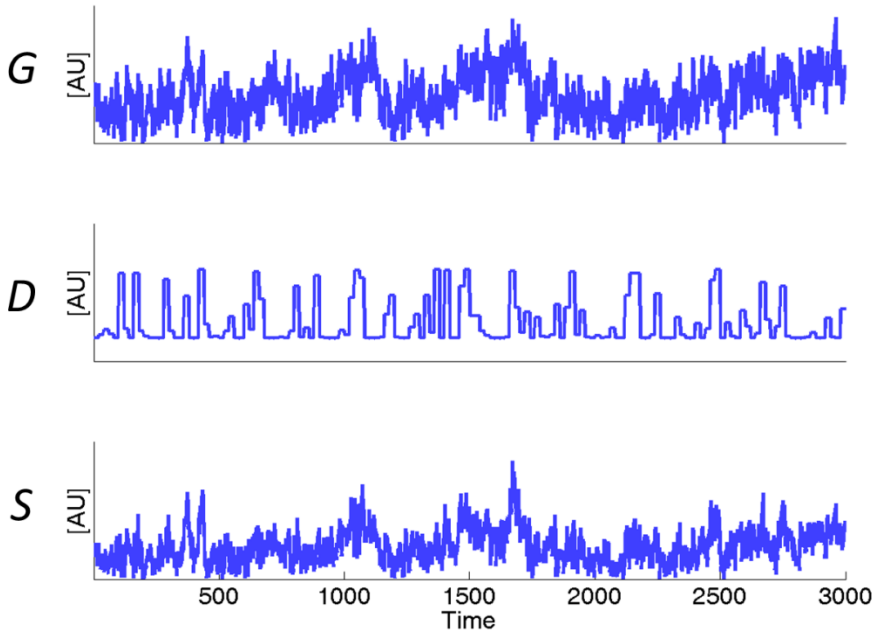

Figure 6.4. Description of the relationship between total neural excitability G^t, sensory drive D^t and the simulated output S^t. In this illustration, the unit prefers a particular motion direction, and stimulus motion directions are uniformly sampled, resulting in the "block-like" fluctuations in D^t. As we see in the bottom row, the presynaptic input S^t is a combination of the noisy excitability fluctuation and true sensory drive.

For each unit i at the PRE layer (number of units $N_{PRE} = 108$), the sensory drive D_i^t is determined by a tuning function of θ^t, which is the stimulus direction. The stimulus θ^t at time t is defined in terms of motion direction, so that $\theta^t \in [0°, 360°)$, for any t. The tuning function follows a von-Mises function:

$$D_i^t(\theta^t) = \exp\left(\kappa_{PRE} \cos\left(\theta^t - p_{PRE,i}\right)\right), \quad \ldots (2)$$

where $\kappa_{PRE} = 16$ controls the selectivity (the larger it is, the narrower the tuning function is), and $p_{PRE,i}$ is the preferred stimulus direction of unit i. The maximum and the minimum values of D_i^t are scaled to be at 30 and 0.9, respectively.

Adopted from Stevenson et al.'s model, we assume neural excitability changes across multiple timescales, such that the total excitability G^t of each unit at time t is defined as follows:

$$G^t = 1 + \sum_{m=1}^{M} g_m^t, \quad \ldots (3)$$

where g_m^t is a small fluctuation component of a particular timescale. M = 10 is the number of different timescales used. Each timescale is defined by a time constant τ_m. These M time constants are evenly spaced in the log-space between 2 and 330000. Each small fluctuation g_m^t is sampled at every time t via

$$g_m^t | g_m^{t-1} \sim N\left(\left(1 - \frac{\Delta t}{\tau_m}\right) g_m^{t-1}, c\left(\frac{\Delta t}{\tau_m}\right)\right), \quad \ldots (4)$$

so that the smaller the τ_m, the faster the dynamical sequence of g_m^t converges back to g_m^0, which is set at 0 for all m, as shown in Figure 6.5. This defines a dynamical system in which each small excitability moves around 0 across time, with different "drifting" speeds determined by τ_m. Δt

indicates the width of each time step, and is fixed as 1. c is a constant that scales the time step, which is fixed as $.01$ for all units and layers.

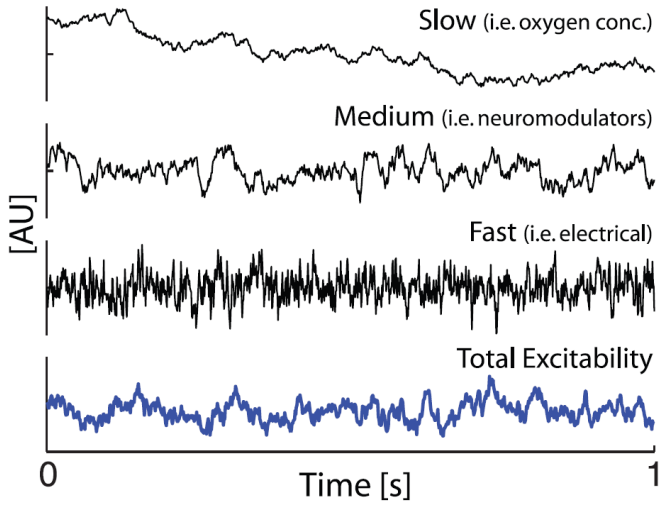

Figure 6.5. Excitability on multiple timescales (adopted from Stevenson et al. (2010)). Excitability of neural units are assumed to fluctuate over multiple timescales, due to various causes (as shown in the figure). The total excitability $\boldsymbol{G^t}$ of a unit at a certain time is assumed to be one plus the sum of the excitabilities across all timescales, so that it can act as a ratio that fluctuates around 1.

From PRE to $V1$

At the $V1$ layer with 72 units (i.e., $N_{V1} = 72$), direction selectivity of each unit j is simulated by different weights assigned to different PRE units. Again, we used the von-Mises function to implement this.

$$w_{ij} = \exp\left(\kappa_{V1} \cos\left(p_{V1,j} - \mathrm{p}_{PRE,i}\right)\right) \quad \ldots (5)$$

Large weight values are assigned between a $V1$ unit and a PRE unit if the two units have similar preferred stimulus direction. This means that each $V1$ unit "listens" most to the PRE unit that has the most similar preferred stimulus direction. The weights are normalized so that $\sum_i w_{ij} = 1$. Similar as κ_{PRE}, $\kappa_{V1} = 12$ controls the selectivity of $V1$ units.

The adaptation of units at $V1$ is modeled as the active, continuous estimation of excitabilities of PRE units. The estimation algorithm is described in detail in the next section. The estimate of excitability $\hat{G}^t_{PRE,i}$ for each PRE unit i is then used to normalize the output from that PRE unit. Each $V1$ unit j then combines the normalized output activities from the PRE level by the direction-tuned weighted sum

$$\hat{R}^t_{V1,j} = \sum_{i=1}^{N_{PRE}} w_{ij} \frac{S^t_i}{\hat{G}^t_{PRE,i}} \quad \ldots (6)$$

Estimation of PRE's excitability by $V1$ units

The computational problem faced by a $V1$ neuron is to estimate $\hat{G}^t_{PRE,i}$ for each PRE unit i given its output S^t_i at every time point t. For simplicity, we drop the PRE and the i notations in this section, as the algorithm is identical for each $PRE \rightarrow V1$ connection. We formulate this problem as a sequential Bayesian estimation problem, which involves two steps at each time point t. Figure 6.6 shows the generative model of how $V1$ units assume neural activities in the PRE layer are generated.

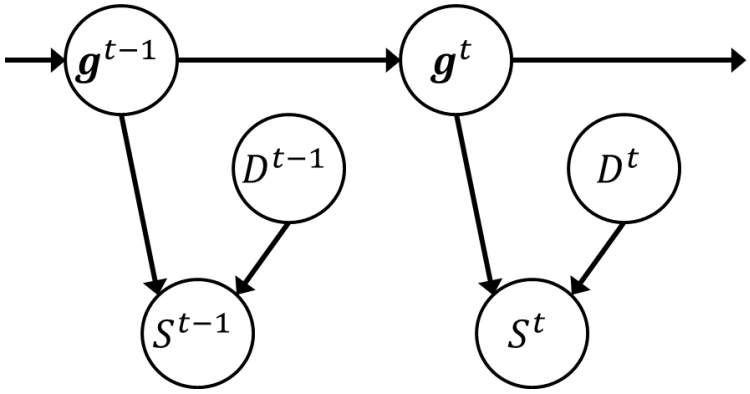

Figure 6.6. The generative model of neural activity from $\boldsymbol{PRE}$ to $\boldsymbol{V1}$. The bold $\boldsymbol{g^t}$ notation indicates a vector containing small components of neural excitabilities that have different timescales.

In the following description, we use $\boldsymbol{g}^t = \{g_m^t\}$ to represent a vector that contains all the small components of g_m^t, for all timescales m. Based on equation (3), the estimate $\hat{G}^t$ can be obtained by estimating $\hat{\boldsymbol{g}}^t$. Given the presynaptic input S^t, the posterior of $\boldsymbol{g}^t$ is given by

$$P(\boldsymbol{g}^t|S^t) \propto P(S^t|\boldsymbol{g}^t)P(\boldsymbol{g}^t|\boldsymbol{g}^{t-1}) \quad \text{... (7)}$$

The temporal prior $P(\boldsymbol{g}^t|\boldsymbol{g}^{t-1})$ is a multivariate normal distribution. We assume the component excitabilities g_m^t are independent. Therefore, $P(\boldsymbol{g}^t|\boldsymbol{g}^{t-1})$ can be simplified as M independent univariate normal distributions described in equation (4).

The likelihood $P(S^t|\boldsymbol{g}^t)$ is written as an integral of D^t as follows

$$P(S^t|\boldsymbol{g}^t) = \int P(S^t|\boldsymbol{g}^t, D^t)P(D^t)\, dD^t \quad \text{... (8)}$$

We implemented the two-step estimation using particle filtering with $Q = 500$ particles. At each time point t, each particle q contains a sample of $\boldsymbol{g}_q^t$. The values of $\boldsymbol{g}_q^t$ in each particle are updated based on the prediction and correction steps described below.

Prediction

In the Prediction step, we implement the temporal dependence of excitability $P(\boldsymbol{g}^t|\boldsymbol{g}^{t-1})$ described in (7). For each particle q, we independently sample $g_{m,q}^t$ based on equation (4)

$$g_{m,q}^t|g_{m,q}^{t-1} \sim N\left(\left(1-\frac{\Delta t}{\tau_m}\right)g_{m,q}^{t-1}, c\left(\frac{\Delta t}{\tau_m}\right)\right) \quad \ldots (9)$$

Each sample of $g_{m,q}^t$ is expected to be close to $g_{m,q}^{t-1}$, and the variance depends on the time constant τ_m.

Correction

In the Correction step, the sampled values of $\boldsymbol{g}_q^t$ are reconciled with the observed S^t via weighted resampling. This is to couple the prior term $P(\boldsymbol{g}^t|\boldsymbol{g}^{t-1})$ in with the likelihood term $P(S^t|\boldsymbol{g}^t)$. In equation (8), we assume, when estimating the excitability of PRE units, $V1$ units hold the assumption that sensory drive is sparse (Stevenson et al., 2010), i.e.

$$P(D^t) = e^{-D^t} \quad \ldots (10)$$

Substituting equation (10) into equation (8) and applying the transformation of $D^t = S^t/G^t$, we obtain

$$P(S^t|\boldsymbol{g}^t) = \frac{1}{G^t}e^{-S^t/G^t} \quad \ldots (11)$$

Based on equation (11), we first compute the likelihood value L_q for each particle based on its $\boldsymbol{g}_q^t$ values. Then, we normalize L_q so that they sum to one and can be used as resampling weights. We resample a new set of Q particles based on their normalized weights as a sampling distribution without replacement, so that particles with larger likelihood values will be resampled more often, and particles with smaller likelihood values less often. We take this set of resampled

particles $\boldsymbol{g}_q^t$ as a representative sample of the posterior of $P(\boldsymbol{g}^t|S^t)$, and obtain an estimate of $\hat{\boldsymbol{g}}^t$ by computing the mean of $\boldsymbol{g}^t$ across all particles:

$$\hat{g}_m^t = \frac{1}{Q}\sum_q g_{q,m}^t \quad \ldots (12)$$

The estimate of the total excitability can then be obtained based on equation (3)

$$\hat{G}^t = 1 + \sum_{m=1}^{M} \hat{g}_m^t \quad \ldots (13)$$

From $V1$ to MT

With the estimated excitabilities $\hat{G}_{PRE,i}^t$ of all PRE units i, the output activities $\hat{R}_{V1,j}^t$ of each $V1$ unit j can then be computed as described in equation (6). These $\hat{R}_{V1,j}^t$ activities are then fed to the units at the MT level as input. Similar to the connections between PRE and $V1$ described in equation (5), the connection between each $V1$ unit j and each MT unit k is defined as

$$w_{jk} = \exp\left(\kappa_{MT}\cos\left(p_{MT,k} - \mathrm{p}_{V1,j}\right)\right) \quad \ldots (14)$$

We normalized the weights w_{jk} so that $\sum_j w_{jk} = 1$. The output of each MT unit k is computed similarly as that of a $V1$ unit described in equation (6)

$$\hat{R}_{MT,k}^t = \sum_{j=1}^{N_{V1}} w_{jk}\frac{\hat{R}_{V1,j}^t}{\hat{G}_{V1,j}^t} \quad \ldots (15)$$

The main computational problem faced by the MT units is to estimate the excitabilities $\hat{G}_{V1,j}^{t}$ of each $V1$ unit j based on its output $\hat{R}_{V1,j}^{t}$. We use the same particle-filtering algorithm described above to compute $\hat{G}_{V1,j}^{t}$.

We postulate that the main difference between the $V1 \rightarrow PRE$ estimation of presynaptic excitabilities and the $MT \rightarrow V1$ estimation of presynaptic excitabilities is the difference in the assumption held by the postsynaptic units about sensory statistics. Instead of assuming sensory drive is temporally-independent, MT units assume sensory drive to be sparse or temporally-smooth (see Figure 6.7). This is in line with a previous study on natural motion statistics (Roth & Black, 2007), which shows that the distribution of temporal derivatives of optical flows has a sharp peak at zero.

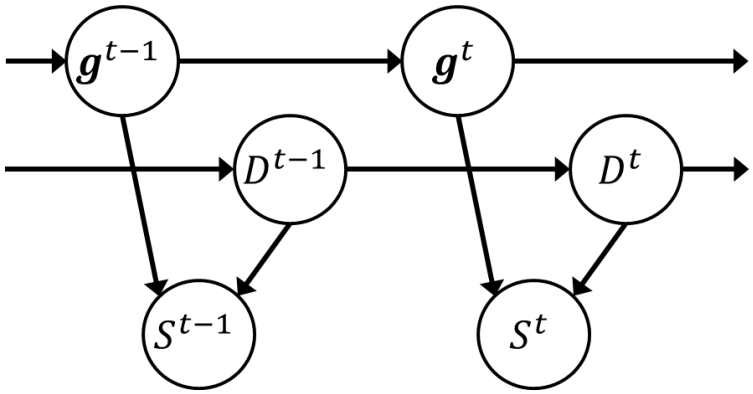

Figure 6.7. The generative model of neural activity from $V1$ to MT. The main difference between this and the $\boldsymbol{V1 \rightarrow PRE}$ model (Figure 6.6) is the temporal dependence between $\boldsymbol{D^{t-1}}$ and $\boldsymbol{D^{t}}$.

Formally, we write the temporal prior for sensory drive as follows

$$P(D^t|D^{t-1}) = \lambda\frac{e^{-\alpha D^t}}{\alpha} + (1-\lambda)\frac{e^{-\beta|D^t - D^{t-1}|}}{Z(D^{t-1},\beta)}$$

... (16)

This prior is a mixture model of motion statistics. The first term on the right-hand side is the "sparseness" term that resembles the function used in the $V1 \rightarrow PRE$ estimation. It is essentially an exponential distribution, which assigns high probabilities to small sensory drive. The second term is the temporal "smoothness" term. The power of the exponent is the negative of the absolute difference between sensory drive across two adjacent time points, $-|D^t - D^{t-1}|$. The smoothness term is defined this way so that, at any time point, it assigns a higher probability to sensory drive that is similar to that of the previous time point. The Z term in the denominator is a normalization term that depends on D^{t-1} and β. The two preferences are weighted by λ, which depends on the magnitude of sensory drive D^{t-1}

$$\lambda = \frac{1}{1 + rD^{t-1}}$$

... (17)

When sensory drive D^{t-1} is high at time t, λ becomes small, so that the prior weights more on the smoothness term than on the sparseness term. When sensory drive D^{t-1} is low at time t, λ becomes large, so that the prior relies more on the sparseness term than on the smoothness term. This is to capture the intuition that when a motion signal is present, it tends to stay on for some time instead of ending abruptly.

There are three free parameters in this temporal prior on sensory drive, namely, α, β and r. The first two parameters, α and β, scale the sparseness and smoothness powers respectively so that they are within comparable range. We fixed $\alpha = 1$ and $\beta = 70$, but model performance is

similar across a reasonable range around these values. The parameter r controls how much the contribution of the two terms depends on the value of sensory drive. When $r = 0$, the model completely relies on the sparseness prior, regardless of the value of D^t. When r gets larger, the contribution of the sparseness term decreases and the prior relies on the smoothness term more.

Results

In our simulation, sensory adaptation was simulated by presenting the same stimulus motion direction to the full model consecutively for at least 500 time steps. Results obtained *before* and *after* the adaptation are then compared. For simplicity, we assume the estimate of each presynaptic unit remains constant from the final time step of adaptation to the test time step. Although the main goal of our simulation is to predict adaptation-induced tuning changes demonstrated by neurons in MT, we also extracted results at each level of simulation to observe the adaptation-induced changes at each processing level. These results are presented in this section as follows. First, at the level of single synapse, we show that the normalization of presynaptic input using excitability estimate can reduce response variability (Stevenson, et al., 2010) which provides a potential functional benefit for adaptation in general. Second, the model predicts typical adaptation-induced changes found in neurons in V1 after motion adaptation. Last, the model qualitatively predicts adaptation-induced changes found in neurons in MT (Kohn & Movshon, 2004).

Single synapse

In order to reliably represent sensory stimulation, the postsynaptic unit should produce output that is as stable as possible if the presynaptic unit is stimulated by the same stimulus within a

short time window (e.g., 50 timesteps). By normalizing presynaptic input using the estimated excitability, large amount of variability in the input can be filtered out to produce a more stable output (red line in Figure 6.8). When compared with the case in which the postsynaptic output is equal to the presynaptic input (blue line in Figure 6.8), the standard deviation in responses can be greatly reduced by at least 60%.

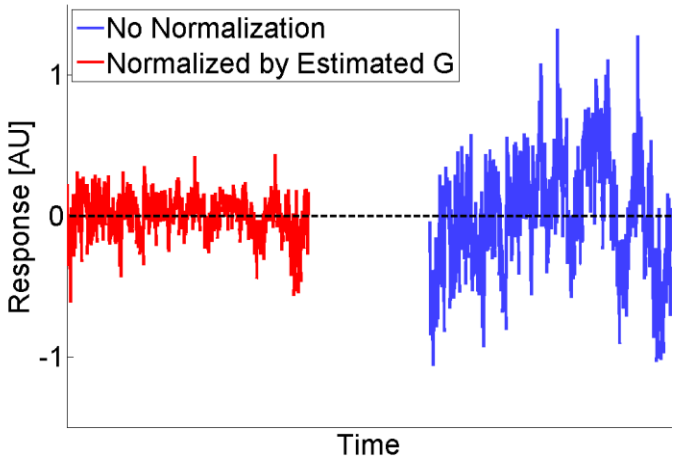

Figure 6.8. Output of a postsynaptic unit when the same stimulus is presented to the presynaptic unit; with vs without normalization by estimated excitability. The red line on the left is the postsynaptic output simulated based on equation (6), with $\boldsymbol{w_{ij} = 1}$ and focusing on only on one presynaptic unit $\boldsymbol{i}$ and one postsynaptic unit $\boldsymbol{j}$. The blue line on the right is simply to produce the presynaptic activity as output without any normalization.

V1 tuning changes

We extracted tuning curves of $V1$ units before and after adaptation. In our simulation, adaptation-induced changes in tuning curves resemble those found in neurons in V1 in primate brain after orientation adaptation (Dragoi, et al., 2000). Figure 6.9 shows the overall changes in tuning across multiple V1 units, which captures the main pattern of response reduction and

broadening of bandwidth at the adapting stimulus. Small repulsive shift of tuning peaks away from the adaptor can also be observed.

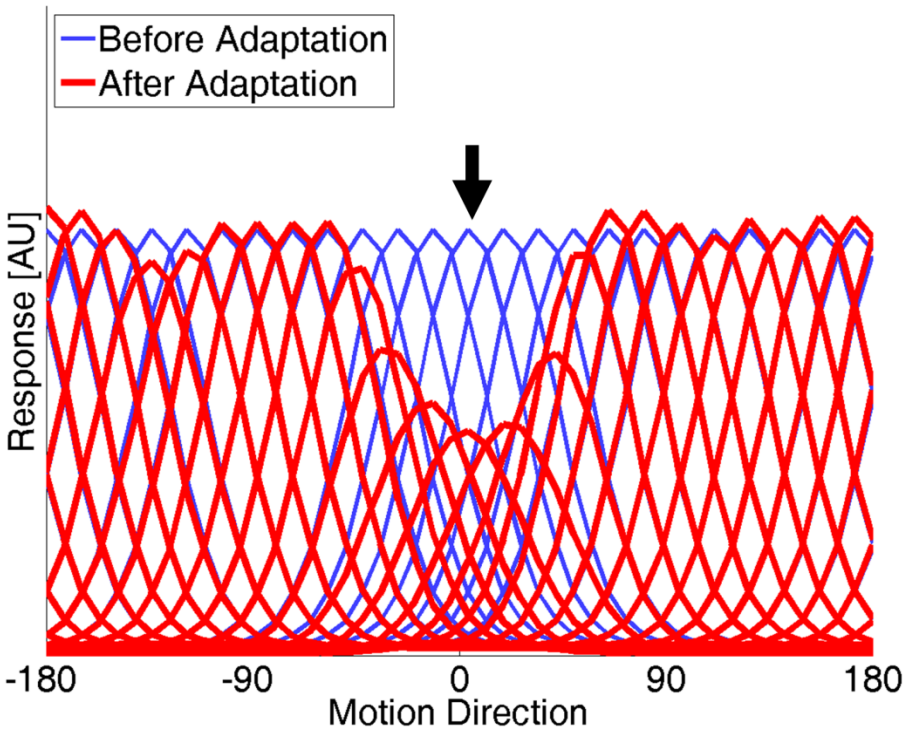

Figure 6.9. Simulated adaptation-induced tuning changes in V1 population. Adapting stimulus was presented at 0°. Only 24 units' tuning curves are shown for clarity.

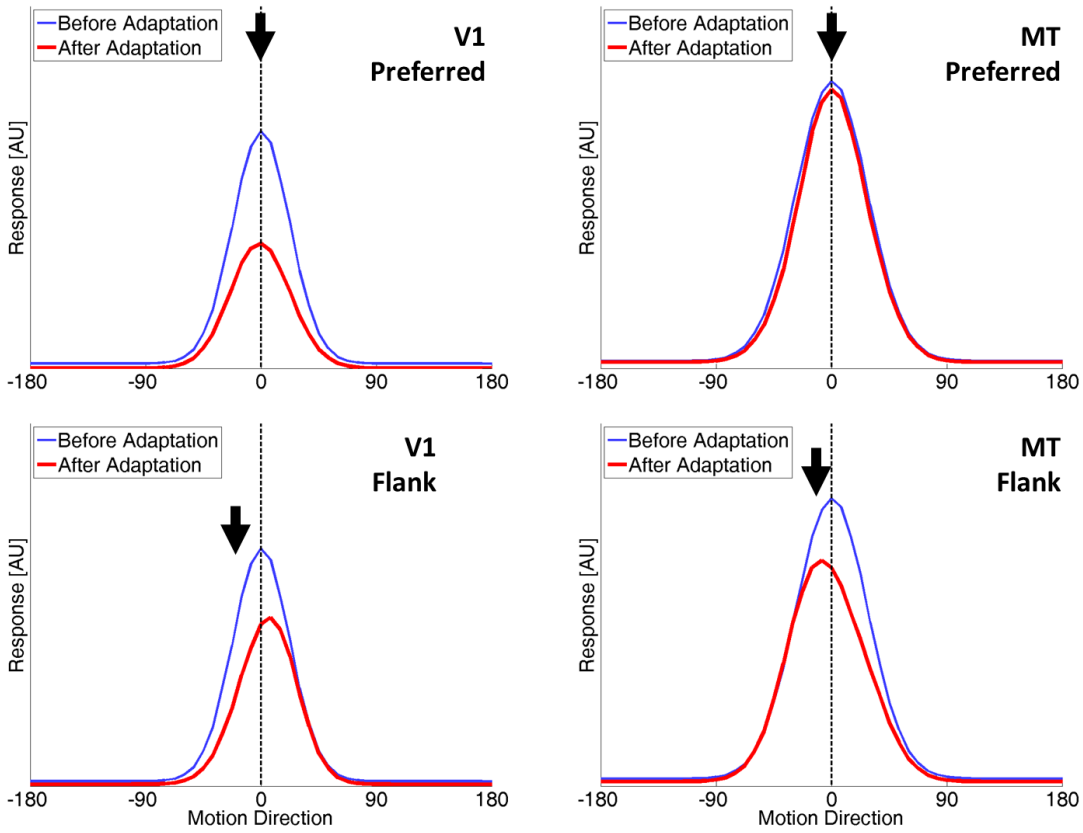

Figure 6.10. Simulated tuning changes at V1 (left two panels) and MT (right two panels) layers after adapting to the unit's preferred (top two panels) and flank (bottom two panels) directions. Adaptation was induced after 500 time steps of presentation of the adaptor. Arrows indicate the adapting direction. Preferred direction of the unit is 0° before adaptation. Blue and red lines plot tuning curves before and after adaptation, respectively.

MT tuning changes

Our simulation can produce adaptation-induced tuning changes at the MT level that are qualitatively similar to those found in neurons in primate MT (Kohn & Movshon, 2004). In particular, MT units in our model can produce the two characteristic changes: 1) the narrowing of tuning bandwidth when the adaptor overlaps with the units preferred direction (Figure 6.10,

top-right); and 2) the attractive shift of tuning peak when the adaptor is on the flank of the unit's tuning curve the preferred direction (Figure 6.10, bottom-right). These changes are plotted in the same figure in contrast with the opposite changes observed in V1 units (Figure 6.10, broadening of tuning bandwidth, top-left; repulsive shift of tuning peak, bottom-left).

General Discussion

We demonstrated that adaptation-induced tuning changes found in V1 and MT can be simultaneously explained in a hierarchical model, with units at different processing levels holding different statistical assumptions about the sensory world. In particular, V1 units in our model demonstrate the typical reduction in responses, repulsive shift of peaks and broadening of tuning. On the other hand, MT units show narrowing and attractive shift of peaks after adaptation. Our simulation results are not only consistent with previous neurophysiological findings (Dragoi, et al., 2000; Kohn & Movshon, 2004), but also related to the pattern of psychophysical results on the direction aftereffect reported in the previous chapter. The stronger DAE observed after global level adaptation is related to the adaptation-induced attractive shift of tuning preference, which is captured in the adapted output of the MT level of our model. Thus, our model explains the results of the previous chapter by explicitly detailing the underlying computational principles that lead to difference in DAE strength between local and global motion adaptation.

Why does the model work?

From PRE to $V1$, our model works based on the same principle as in Stevenson et al.'s (2010) model. During adaptation, the same adapting stimulus direction is presented over time. This leads to persistently highly sensory drive D^t sent to the presynaptic units whose preferred directions are close to the adapting direction. This causes the presynaptic activity S^t to maintain

at a high level during adaptation. Because, at this level, the postsynaptic neuron assumes sensory drive is sparse, the high input activity S^t is thought to be unlikely to be caused by persistently high sensory drive. Instead, the postsynaptic neuron attributes the high input activity to high presynaptic excitability G^t. This leads to an overestimation of presynaptic excitability $\hat{G}_i^t$. Because the $\hat{G}_i^t$ for presynaptic unit i is placed at the denominator as a normalizing constant, the larger the presynaptic input S_i^t, the greater the suppression it takes due to the overestimation of $\hat{G}^t$. In short, the suppression of input is the greatest for units with persistently high presynaptic activities. This property nicely agrees with the phenomenon of adaptation-induced response gain reduction, and, therefore, can capture many low-level adaptation phenomena (Stevenson, et al., 2010).

However, from $V1$ to MT, the model works differently. Both the sparseness term and the temporal-smoothness term in the prior affect the estimation of presynaptic excitability. The sparseness term, in general, works in exactly the same way as the sparseness assumption made by $V1$ units, which states that sensory drive is sparse. When it kicks in, the sparseness term increases excitability estimates, and suppresses the presynaptic units that fire the most as described by above. However, the smoothness term assigns higher probabilities to sensory drive values that have small difference over time. As a result, when faced with persistently high input, the smoothness term allows an MT unit to attribute it to high sensory drive instead of high excitability. This results in an underestimation of excitabilities in the $V1$ units, producing an enhancement of output.

The critical part of the mixture prior is the weighted-sum that combines the sparseness and the smoothness terms. The weight is determined by the magnitude of the sensory drive, which is represented in the presynaptic activity. Although the output activities from the $V1$ units have

already been adapted (normalized) and reduced, those units that prefer a motion direction close to the adaptor still produce persistently higher activities than others. From the perspective of a MT unit, persistently high presynaptic input reduces the weight (λ) for the sparseness term, but increases the weight $(1-\lambda)$ for the smoothness term. This way, the smoothness term kicks in and the sparseness term is suppressed, so that model attributes the persistently high input to temporally smooth sensory drive. Thus, it regards high presynaptic excitability as an unlikely cause of the high input, and underestimates the excitability $\hat{G}_{V1}^{t}$. As a corollary of the above argument, smaller $\hat{G}_{V1}^{t}$ values lead to higher output from the MT unit, which leads to specific response facilitation *around* the adapting direction. Such facilitation underlies the observed neural adaptation phenomena of narrowing of tuning and attractive shift of peak towards the adaptor.

Limitations

Despite its power in predicting different types of adaptation-induced tuning changes, we note that the model proposed in this chapter has certain limitations. They can all be summarized as issues about the physiological plausibility of the model structure. First, the model is insensitive to the location of stimulus presentation. As motion-sensitive neurons usually have receptive fields confined to certain locations or regions, it becomes a problem for the model when predicting adaptation-induced changes that are location-specific (Kohn & Movshon, 2003). This also limits its ability to predict perceptual aftereffects that are due to manipulation of adapting and test locations, such as those reported in previous chapters of this dissertation (Lee & Lu, 2012), and in some other studies (e.g., Snowden & Milne, 1997). Second, the model is insensitive to stimulus size, which has recently been found to have great effect in adaptation-induced changes in tuning of V1 neurons (Wissig & Kohn, 2012). Third, the model has a

completely feedforward architecture, and does not contain lateral and feedback connections. These connections exist and are crucial in explaining phenomena related to surround suppression and contextual effects. Indeed, the model proposed in this chapter is a simple model in terms of the architecture, and may not be readily generalized to explain a wide range of neural and perceptual consequence of adaptation. Nevertheless, the power in predicting level-specific adaptation-induced neural changes grants the framework a promising starting point for future research.

Relevance to previous models

Many previous models have used priors that describe natural sensory statistics in a Bayesian framework in modeling visual perception (Weiss, et al., 2002; Yuille & Grzywacz, 1988). Stevenson et al.'s model is the first one to incorporate natural sensory statistics (sparseness) in modeling adaptation. In our model, we push the idea further to incorporate higher-order natural sensory statistics (temporal smoothness of motion) to predict level-specific neural changes caused by adaptation. The idea that high-level processing may employ more complex prior assumptions is supported by the findings in the first chapter of this dissertation (Lee & Lu, 2010), as well as some other computational study (Wu, et al., 2009). Moreover, this idea is related to the predictive coding framework (Rao & Ballard, 1999) in that what high-level units assume is what exists in the natural environment, which can be understood as a "default prediction" of stimulus in the external world. Indeed, similar ideas of how natural statistics may have shaped the sensory system have been demonstrated (Olshausen & Field, 1996), discussed (Schwartz, et al., 2007) and applied to model neural and perceptual effects of spatial contextual (Schwartz, Sejnowski, & Dayan, 2009).

Future directions

As mentioned in the previous paragraph, what our model is not yet able to explain are some perceptual aftereffects that are related to how sensory neurons are connected in a circuitry. One possible direction for future research along this direction is to incorporate natural sensory statistics into a more physiologically-plausible structure in order to explain a wider range of neural and perceptual adaptation phenomena.

Chapter 7

Conclusion

Summary and Conclusion

Can we study hierarchical adaptation at individual levels of processing? Findings from Chapters 2 and 3 give a positive answer to this question. Findings in Chapter 2 suggest that integration of complex motion patterns may differ from that of simple translational patterns, implying the existence of some higher-order integration mechanisms. Results of experiments in Chapter 3 suggest that there exist at least two levels of motion adaptation, which can be separated probed using psychophysical methods. Taken together, these results support the notion that the hierarchical motion system does not adapt as a whole, and individual levels of adaptation can be separately investigated.

If one could specifically study the effects of adaptation at individual processing levels, how does the adaptation of each level contribute in producing perceptual aftereffects? Chapter 4 shows that adaptation of the local motion processing level *per se* is sufficient in producing a MAE when their effects are propagated downstream. In Chapter 5, by independently adapting the local and the global levels of motion processing, I compared the strength of direction aftereffect based on these two levels of adaptation, and found that global adaptation leads to stronger direction aftereffects. Results from these two chapters suggest that 1) adaptation of low and high levels of motion processing may have different roles to play in generating perceptual aftereffects, and 2) propagation of low-level adaptation effects may involve spatial integration of local illusory signals due to local adaptation.

Finally, in Chapter 6, I investigated the computational principle that underlies different neural adaptation phenomena observed at different motion processing levels. I attempted to explain different adaptation-induced tuning changes in V1 and MT are caused by different assumptions held by neurons about the natural sensory environment. I have shown that a sparse-or-smooth prior on sensory drive can explain the characteristic changes in neural tuning found in MT neurons caused by motion adaptation.

Taken together, findings from this dissertation illustrate that adaptation of the sensory system should be investigated in the context of hierarchy processing. Especially for the motion system, perceptual aftereffects depend on the adaptation of both local and global levels of processing. Difference in neural adaptation at the local and global motion levels can be explained from a natural-statistics perceptive: neurons at different levels are holding different assumptions about natural sensory statistics. In conclusion, when studying sensory adaptation, one should not only focus at one level of processing. Instead, investigating how different levels adapt, and how adaptation effects at one level might affect the processing of the other may shed more lights on the big picture of hierarchical adaptation.

General Discussion

There are at least three limitations regarding the findings reported in this dissertation. First, in Chapters 4 and 5, perceptual aftereffects generated by local and global levels of adaptation were directly compared in terms of strength. One may be concerned whether such comparison is a fair one. Because, in the Local conditions, elements were equally divided into five sets moving in different global directions, the Global conditions had five times more coherent elements than that in the Local conditions. Difference in aftereffect strength, particularly the Global > Local trend

reported in those experiments, may be caused by the larger number of coherently-moving adapting elements in the Global condition. This is a legitimate concern, but results from the density experiment show that global direction discrimination did not change much even with a ten-fold increase/decrease of stimulus density. Given such pattern of results, a more conservative way of testing the "Global > Local" trend may be to use just one-fifth of the elements in the adapting phase in the Global conditions, in order to match with the number of coherently-moving elements in the Local conditions. If the same trend persisted even with this manipulation, it might provide a clearer picture about the comparison of aftereffect strengths across the two levels of adaptation.

Second, all experiments were conducted using the multiple-aperture stimulus. As mentioned in individual chapters, it is a good tool for probing different levels of the motion system for the effects of adaptation. However, it is unclear if the psychophysical results observed could be generalized to other types of stimuli. For example, one could use a similar type of stimulus with multiple-aperture over a real scene of motion sequence (Kane, Bex, & Dakin, 2011) as the adapting stimulus. There may exist some top-down influence (Lin & He, 2012; Scarfe & Johnston, 2011; Winawer, Huk, & Boroditsky, 2008) from the meaning of the scene and other aspects, such as attention (Huk, Ress, & Heeger, 2001), that could affect perceptual aftereffects. Especially in conditions in which the adapting patterns is by no means a natural or realistic scene (e.g., the 5-directional patterns in Chapters 4 and 5), it is possible that what was observed was only a special case of how motion adaptation could happen in the human visual system.

Third, none of the experiments reported here address the issue of the timescales for perceptual aftereffects generated by different levels of adaptation. For example, the local and global levels of adaptation could have produced different strengths of aftereffects at different points. This is

particularly relevant to experiments in which the test duration was long (e.g., Chapter 3). While it was not addressed in any experiments in this dissertation, the experiments provide a good starting point for one to tease out the temporal dynamics of different levels of adaptation. For example, one could employ the 2x2 design in Chapter 5 but measure some duration-related aspects of aftereffects, e.g., duration of the aftereffect, how adapting duration may affect aftereffect strength, etc. This would allow one to investigate how the dynamics of different levels of adaptation may contribute in generating the perceptual aftereffects over time.

On the surface, results from Chapters 4 and 5 are inconsistent: Although, in Chapter 4, there was no quantitative measure of how strong perceptual effect was, it appeared to be a strong effect just by how clear the results are. Note that the conditions in Chapter 4 are similar to the Local condition in Chapter 5. Given such results in Chapter 4, the inconsistency appeared to be that the strength of aftereffects measured in the Local condition in Chapter 5 was much lower than that in the Global condition. One possible interpretation is the static or dynamic nature of the test elements. It is believed that static test stimuli tend to probe more low-level mechanisms, while dynamic ones probe more high-level (or, in fact, both low- and high-level) mechanisms (Mather, et al., 2008). Since all results in Chapter 4 were obtained using static test, and those obtained in Chapter 5 were dynamic, the apparent discrepancy between the two chapters might be related to this difference in test stimulus.

Future Directions

The issues mentioned in the limitations may open up new future directions along the line of hierarchical adaptation. In addition to those mentioned above, some other possible directions are as follows.

What is the functional benefit of having adaptation-induced perceptual aftereffects? In the context of hierarchical adaptation, this question can be interpreted as "How far can effects of adaptation be propagated?" This may require one to go beyond the sensory system, and investigate whether sensory adaptation alters other functions. For example, does the adaptation-induced bias in perception affect subsequent actions? A specific example might be to investigate whether motion aftereffect can bias subsequent visual-guided motor responses. If motor responses could be biased by visual aftereffects, it would provide interesting evidence to support that perceptual aftereffect is not merely an epiphenomenon, but serves some functional purpose. Because the general paradigm used in this dissertation to tease out the two levels of adaptation is novel, it would be interesting to transfer the technique to other areas. For example, using multiple-electrode recording on primate's or cat's brain during and after adaptation, one can then investigate how population of neurons changes their response characteristics after adaptation. This will be informative to the understanding of how adaptation may change the brain, especially about how information is represented in the network.

Finally, future modeling work in other domains of sensory adaptation may attempt to incorporate natural statistics to predicting adaptation-related neural behaviors. When viewed as temporal contextual effects (Schwartz, et al., 2007), adaptation can possibly be studied with other spatial contextual phenomena. It would be interesting to combine spatial (Schwartz, et al., 2009) and temporal contextual models that are relate to natural sensory statistics, in order to provide a more general and unified computational framework for contextual effects.

References

Aaen-Stockdale, C., Ledgeway, T., & Hess, R. F. (2007). Second-order optic flow processing. *Vision Research, 47*(13), 1798-1808.

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal Energy Models for the Perception of Motion. *Journal of the Optical Society of America a-Optics Image Science and Vision, 2*(2), 284-299.

Ahlstrom, U., & Borjesson, E. (1996). Segregation of motion structure from random visual noise. *Perception, 25*(3), 279-291.

Alais, D., Verstraten, F. A., & Burr, D. C. (2005). The motion aftereffect of transparent motion: two temporal channels account for perceived direction. *Vision Res, 45*(4), 403-412.

Amano, K., Edwards, M., Badcock, D. R., & Nishida, S. (2009a). Adaptive pooling of visual motion signals by the human visual system revealed with a novel multi-element stimulus. *J Vis, 9*(3), 4 1-25.

Amano, K., Edwards, M., Badcock, D. R., & Nishida, S. (2009b). Adaptive pooling of visual motion signals by the human visual system revealed with a novel multi-element stimulus. *Journal of Vision, 9*(3(4)), 1-25.

Amano, K., Takeda, T., Haji, T., Terao, M., Maruya, K., Matsumoto, K., et al. (2012). Human neural responses involved in spatial pooling of locally ambiguous motion signals. *J Neurophysiol, 107*(12), 3493-3508.

Anstis, S., Verstraten, F. A. J., & Mather, G. (1998). The motion aftereffect. *Trends in Cognitive Sciences, 2*(3), 111-117.

Baccus, S. A., & Meister, M. (2004). Retina versus cortex: Contrast adaptation in parallel visual pathways. *Neuron, 42*(1), 5-7.

Badcock, D. R., & Khuu, S. K. (2001). Independent first and second-order motion energy analyses of optic flow. *Psychological Research-Psychologische Forschung, 65*(1), 50-56.

Bair, W., & Movshon, J. A. (2004). Adaptive temporal integration of motion in direction-selective neurons in macaque visual cortex. *Journal of Neuroscience, 24*(33), 7305-7323.

Barlow, H., & Tripathy, S. P. (1997). Correspondence noise and signal pooling in the detection of coherent visual motion. *Journal of Neuroscience, 17*(20), 7954-7966.

Barlow, H. B. (1990). A theory about the functional role and synaptic mechanisms of visual aftereffects. In C. Blakemore (Ed.), *Vision: Coding and efficiency* (pp. 363-375). New York: Cambridge University Press.

Barlow, H. B., & Tripathy, S. P. (1997). Correspondence noise and signal pooling in the detection of coherent visual motion. *Journal of Neuroscience, 17*(20), 7954-7966.

Beardsley, S. A., & Vaina, L. A. (2005). Psychophysical evidence for a radial motion bias in complex motion discrimination. *Vision Research, 45*(12), 1569-1586.

Beardsley, S. A., Ward, R. L., & Vaina, L. M. (2003). A neural network model of spiral-planar motion tuning in MSTd. *Vision Research, 43*(5), 577-595.

Bertone, A., & Faubert, J. (2003). How is complex second-order motion processed? *Vision Research, 43*(25), 2591-2601.

Bex, P. J., Metha, A. B., & Makous, W. (1999). Enhanced motion aftereffect for complex motions. *Vision Research, 39*(13), 2229-2238.

Blake, R., & Aiba, T. S. (1998). Detection and discrimination of optical flow components. *Japanese Psychological Research, 40*(1), 19-30.

Blake, R., Tadin, D., Sobel, K. V., Raissian, T. A., & Chong, S. C. (2006). Strength of early visual adaptation depends on visual awareness. *Proc Natl Acad Sci U S A, 103*(12), 4783-4788.

Braddick, O. (1993). Segmentation Versus Integration in Visual-Motion Processing. *Trends in Neurosciences, 16*(7), 263-268.

Braddick, O. J., Wishart, K. A., & Curran, W. (2002). Directional performance in motion transparency. *Vision Res, 42*(10), 1237-1248.

Bradley, D. C., Chang, G. C., & Andersen, R. A. (1998). Encoding of three-dimensional structure-from-motion by primate area MT neurons. *Nature, 392*(6677), 714-717.

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10*(4), 433-436.

Britten, K. H., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1993). Responses of neurons in macaque MT to stochastic motion signals. *Vis Neurosci, 10*(6), 1157-1169.

Burr, D. C., Badcock, D. R., & Ross, J. (2001). Cardinal axes for radial and circular motion, revealed by summation and by masking. *Vision Research, 41*(4), 473-481.

Burr, D. C., Morrone, M. C., & Vaina, L. M. (1998). Large receptive fields for optic flow detection in humans. *Vision Research, 38*(12), 1731-1743.

Burr, D. C., & Santoro, L. (2001). Temporal integration of optic flow, measured by contrast and coherence thresholds. *Vision Research, 41*(15), 1891-1899.

Carandini, M., & Heeger, D. J. (2012). Normalization as a canonical neural computation. *Nature Reviews Neuroscience, 13*(1), 51-62.

Clark, A. M., & Bradley, D. C. (2008). *Integration of distributed one-dimensional motion signals by macaque middle temporal cortical neurons.* Paper presented at the Computational and Systems Neuroscience (Cosyne), Salt Lake City, UT.

Clifford, C. W., Webster, M. A., Stanley, G. B., Stocker, A. A., Kohn, A., Sharpee, T. O., et al. (2007). Visual adaptation: neural, psychological and computational aspects. *Vision Res, 47*(25), 3125-3131.

Clifford, C. W., Wenderoth, P., & Spehar, B. (2000). A functional angle on some after-effects in cortical vision. *Proc Biol Sci, 267*(1454), 1705-1710.

Clifford, C. W. G., Beardsley, S. A., & Vaina, L. M. (1999). The perception and discrimination of speed in complex motion. *Vision Research, 39*(13), 2213-2227.

Crick, F., & Koch, C. (1995). Are We Aware of Neural Activity in Primary Visual-Cortex. *Nature, 375*(6527), 121-123.

Curran, W., Clifford, C. W. G., & Benton, C. P. (2006). The direction aftereffect is driven by adaptation of local motion detectors. *Vision Research, 46*(25), 4270-4278.

Curran, W., Hibbard, P. B., & Johnston, A. (2007). The visual processing of motion-defined transparency. *Proc Biol Sci, 274*(1613), 1049-1056.

Dragoi, V., Sharma, J., & Sur, M. (2000). Adaptation-induced plasticity of orientation tuning in adult visual cortex. *Neuron, 28*(1), 287-298.

Duffy, C. J., & Wurtz, R. H. (1991a). Sensitivity of Mst Neurons to Optic Flow Stimuli .1. A Continuum of Response Selectivity to Large-Field Stimuli. *Journal of Neurophysiology, 65*(6), 1329-1345.

Duffy, C. J., & Wurtz, R. H. (1991b). Sensitivity of Mst Neurons to Optic Flow Stimuli .2. Mechanisms of Response Selectivity Revealed by Small-Field Stimuli. *Journal of Neurophysiology, 65*(6), 1346-1359.

Duffy, C. J., & Wurtz, R. H. (1997). Medial superior temporal area neurons respond to speed patterns in optic flow. *Journal of Neuroscience, 17*(8), 2839-2851.

Dumoulin, S. O., & Wandell, B. A. (2008). Population receptive field estimates in human visual cortex. *Neuroimage, 39*(2), 647-660.

Edwards, M., & Badcock, D. R. (1994). Global Motion Perception - Interaction of the on and Off Pathways. *Vision Research, 34*(21), 2849-2858.

Edwards, M., & Badcock, D. R. (1995). Global Motion Perception - No Interaction between the First-Order and 2nd-Order Motion Pathways. *Vision Research, 35*(18), 2589-2602.

Edwards, M., & Badcock, D. R. (1996). Global-motion perception: Interaction of chromatic and luminance signals. *Vision Research, 36*(16), 2423-2431.

Edwards, M., & Greenwood, J. A. (2005). The perception of motion transparency: a signal-to-noise limit. *Vision Res, 45*(14), 1877-1884.

Emerson, R. C., Bergen, J. R., & Adelson, E. H. (1992). Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Res, 32*(2), 203-218.

Freeman, T. C. A., & Harris, M. G. (1992). Human Sensitivity to Expanding and Rotating Motion - Effects of Complementary Masking and Directional Structure. *Vision Research, 32*(1), 81-87.

Gibson, J. J., & Radner, M. (1937). Adaptation, after-effect and contrast in the perception of tilted lines. I. Quantitative studies. *Journal of Experimental Psychology, 20*, 453-467.

Gilbert, C. D., & Wiesel, T. N. (1990). The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat. *Vision Res, 30*(11), 1689-1701.

Greenwood, J. A., & Edwards, M. (2006a). An extension of the transparent-motion detection limit using speed-tuned global-motion systems. *Vision Research, 46*(8-9), 1440-1449.

Greenwood, J. A., & Edwards, M. (2006b). Pushing the limits of transparent-motion detection with binocular disparity. *Vision Research, 46*(16), 2615-2624.

Greenwood, J. A., & Edwards, M. (2009). The detection of multiple global directions: Capacity limits with spatially segregated and transparent-motion signals. *Journal of Vision, 9*(1).

Grunewald, A., & Lankheet, M. J. (1996). Orthogonal motion after-effect illusion predicted by a model of cortical motion processing. *Nature, 384*(6607), 358-360.

Hammond, P., Pomfrett, C. J., & Ahmed, B. (1989). Neural motion after-effects in the cat's striate cortex: orientation selectivity. *Vision Res, 29*(12), 1671-1683.

He, S., Cavanagh, P., & Intriligator, J. (1996). Attentional resolution and the locus of visual awareness. *Nature, 383*(6598), 334-337.

He, S., & MacLeod, D. I. (2001). Orientation-selective adaptation and tilt after-effect from invisible patterns. *Nature, 411*(6836), 473-476.

Heeger, D. J. (1992). Normalization of Cell Responses in Cat Striate Cortex. *Visual Neuroscience, 9*(2), 181-197.

Heeger, D. J., Simoncelli, E. P., & Movshon, J. A. (1996). Computational models of cortical visual processing. *Proc Natl Acad Sci U S A, 93*(2), 623-627.

Hershenson, M. (1993). Linear and Rotation Motion Aftereffects as a Function of Inspection Duration. *Vision Research, 33*(14), 1913-1919.

Huk, A. C., & Heeger, D. J. (2002). Pattern-motion responses in human visual cortex. *Nature Neuroscience, 5*(1), 72-75.

Huk, A. C., Ress, D., & Heeger, D. J. (2001). Neuronal basis of the motion aftereffect reconsidered. *Neuron, 32*(1), 161-172.

Ibbotson, M. R., & Price, N. S. (2001). Spatiotemporal tuning of directional neurons in mammalian and avian pretectum: a comparison of physiological properties. *J Neurophysiol, 86*(5), 2621-2624.

Jazayeri, M., & Movshon, J. A. (2006). Optimal representation of sensory information by neural populations. *Nature Neuroscience, 9*(5), 690-696.

Jin, D. Z., Dragoi, V., Sur, M., & Seung, H. S. (2005). Tilt aftereffect and adaptation-induced changes in orientation tuning in visual cortex. *Journal of Neurophysiology, 94*(6), 4038-4050.

Kane, D., Bex, P., & Dakin, S. (2011). Quantifying "the aperture problem" for judgments of motion direction in natural scenes. *Journal of Vision, 11*(3).

Kanwisher, N. (2001). Neural events and perceptual awareness. *Cognition, 79*(1-2), 89-113.

Kohn, A. (2007). Visual adaptation: Physiology, mechanisms, and functional benefits. *Journal of Neurophysiology, 97*(5), 3155-3164.

Kohn, A., & Movshon, J. A. (2003). Neuronal adaptation to visual motion in area MT of the macaque. *Neuron, 39*(4), 681-691.

Kohn, A., & Movshon, J. A. (2004). Adaptation changes the direction tuning of macaque MT neurons. *Nature Neuroscience, 7*(7), 764-772.

Kontsevich, L. L., & Tyler, C. W. (1999). Bayesian adaptive estimation of psychometric slope and threshold. *Vision Res, 39*(16), 2729-2737.

Krekelberg, B., Dannenberg, S., Hoffmann, K. P., Bremmer, F., & Ross, J. (2003). Neural correlates of implied motion. *Nature, 424*(6949), 674-677.

Krekelberg, B., van Wezel, R. J. A., & Albright, T. D. (2006). Adaptation in macaque MT reduces perceived speed and improves speed discrimination. *Journal of Neurophysiology, 95*(1), 255-270.

Lee, A. L. F., & Lu, H. (2010). A comparison of global motion perception using a multiple-aperture stimulus. *Journal of Vision, 10*(4).

Lee, A. L. F., & Lu, H. (2012). Two forms of aftereffects induced by transparent motion reveal multilevel adaptation. *Journal of Vision, 12*(4).

Lee, H. A., & Lee, S. H. (2012). Hierarchy of direction-tuned motion adaptation in human visual cortex. *Journal of Neurophysiology, 107*(8), 2163-2184.

Levinson, E., & Sekuler, R. (1976). Adaptation Alters Perceived Direction of Motion. *Vision Research, 16*(7), 779-&.

Lin, Z. C., & He, S. (2012). Emergent Filling In Induced by Motion Integration Reveals a High-Level Mechanism in Filling In. *Psychological Science, 23*(12), 1534-1541.

Lochmann, T., Ernst, U. A., & Deneve, S. (2012). Perceptual Inference Predicts Contextual Modulations of Sensory Responses. *Journal of Neuroscience, 32*(12), 4179-4195.

Lopez-Moliner, J., Smeets, J. B. J., & Brenner, E. (2004). Components of motion perception revealed: two different after-effects from a single moving object. *Vision Research, 44*(22), 2545-2549.

Lu, H., Lin, T., Lee, A., Vese, L., & Yuille, A. L. (2010). Functional form of motion priors in human motion perception. In J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. S. Zemel & A. Culotta (Eds.), *Advances in neural information processing systems* (Vol. 23, pp. 1495-1503). Cambridge, MA: MIT Press.

Lu, H., & Yuille, A. L. (2006). Ideal observers for detecting motion: Correspondence noise. In B. Schölkopf, J. Platt & T. Hofmann (Eds.), *Advances in neural information processing systems* (Vol. 18, pp. 827-834). Cambridge, MA: MIT Press.

Marr, D., & Ullman, S. (1981). Directional selectivity and its use in early visual processing. *Proc R Soc Lond B Biol Sci, 211*(1183), 151-180.

Marshak, W., & Sekuler, R. (1979). Mutual Repulsion between Moving Visual Targets. *Science, 205*(4413), 1399-1401.

Maruya, K., Watanabe, H., & Watanabe, M. (2008). Adaptation to invisible motion results in low-level but not high-level aftereffects. *J Vis, 8*(11), 7 1-11.

Mather, G. (1980). The Movement Aftereffect and a Distribution-Shift Model for Coding the Direction of Visual Movement. *Perception, 9*(4), 379-392.

Mather, G., & Moulden, B. (1980). A Simultaneous Shift in Apparent Direction - Further Evidence for a Distribution-Shift Model of Direction Coding. *Quarterly Journal of Experimental Psychology, 32*(May), 325-333.

Mather, G., Pavan, A., Campana, G., & Casco, C. (2008). The motion aftereffect reloaded. *Trends in Cognitive Sciences, 12*(12), 481-487.

Maunsell, J. H. R., & Vanessen, D. C. (1983). Functional-Properties of Neurons in Middle Temporal Visual Area of the Macaque Monkey .1. Selectivity for Stimulus Direction, Speed, and Orientation. *Journal of Neurophysiology, 49*(5), 1127-1147.

McGraw, P. V., Whitaker, D., Skillen, J., & Chung, S. T. (2002). Motion adaptation distorts perceived visual position. *Curr Biol, 12*(23), 2042-2047.

Mingolla, E., Todd, J. T., & Norman, J. F. (1992). The Perception of Globally Coherent Motion. *Vision Research, 32*(6), 1015-1031.

Morrone, M. C., Burr, D. C., & Vaina, L. M. (1995). 2 Stages of Visual Processing for Radial and Circular Motion. *Nature, 376*(6540), 507-509.

Morrone, M. C., Tosetti, M., Montanaro, D., Fiorentini, A., Cioni, G., & Burr, D. C. (2000). A cortical area that responds specifically to optic flow, revealed by fMRI. *Nature Neuroscience, 3*(12), 1322-1328.

Muller, J. R., Metha, A. B., Krauskopf, J., & Lennie, P. (1999). Rapid adaptation in visual cortex to the structure of images. *Science, 285*(5432), 1405-1408.

Nishida, S., & Ashida, H. (2001). A motion aftereffect seen more strongly by the non-adapted eye: evidence of multistage adaptation in visual motion processing. *Vision Res, 41*(5), 561-570.

Olshausen, B. A., & Field, D. J. (1996). Emergence of simple-cell receptive field properties by learning a sparse code for natural images. *Nature, 381*(6583), 607-609.

Orban, G. A., Lagae, L., Raiguel, S., Xiao, D. K., & Maes, H. (1995). The Speed Tuning of Medial Superior Temporal (Mst) Cell Responses to Optic-Flow Components. *Perception, 24*(3), 269-285.

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437-442.

Price, N. S. C., Greenwood, J. A., & Ibbotson, M. R. (2004). Tuning properties of radial phantom motion aftereffects. *Vision Research, 44*(17), 1971-1979.

Prins, N., & Kingdom, F. A. A. (2009). Palamedes: Matlab routines for analyzing psychophysical data.

Qian, N., & Andersen, R. A. (1994). Transparent Motion Perception as Detection of Unbalanced Motion Signals .2. Physiology. *Journal of Neuroscience, 14*(12), 7367-7380.

Qian, N., Andersen, R. A., & Adelson, E. H. (1994a). Transparent Motion Perception as Detection of Unbalanced Motion Signals .1. Psychophysics. *Journal of Neuroscience, 14*(12), 7357-7366.

Qian, N., Andersen, R. A., & Adelson, E. H. (1994b). Transparent Motion Perception as Detection of Unbalanced Motion Signals .3. Modeling. *Journal of Neuroscience, 14*(12), 7381-7392.

Rajimehr, R., Vaziri-Pashkam, M., Afraz, S. R., & Esteky, H. (2004). Adaptation to apparent motion in crowding condition. *Vision Res, 44*(9), 925-931.

Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: a functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience, 2*(1), 79-87.

Ross, J., Badcock, D. R., & Hayes, A. (2000). Coherent global motion in the absence of coherent velocity signals. *Current Biology, 10*(11), 679-682.

Roth, S., & Black, M. J. (2007). On the spatial statistics of optical flow. *International Journal of Computer Vision, 74*(1), 33-50.

Rust, N. C., Mante, V., Simoncelli, E. P., & Movshon, J. A. (2006). How MT cells analyze the motion of visual patterns. *Nature Neuroscience, 9*(11), 1421-1431.

Sakata, H., Shibutani, H., Ito, Y., Tsurugai, K., Mine, S., & Kusunoki, M. (1994). Functional-Properties of Rotation-Sensitive Neurons in the Posterior Parietal Association Cortex of the Monkey. *Experimental Brain Research, 101*(2), 183-202.

Scarfe, P., & Johnston, A. (2011). Global motion coherence can influence the representation of ambiguous local motion. *Journal of Vision, 11*(12).

Sceniak, M. P., Ringach, D. L., Hawken, M. J., & Shapley, R. (1999). Contrast's effect on spatial summation by macaque V1 neurons. *Nature Neuroscience, 2*(8), 733-739.

Schrater, P. R., & Simoncelli, E. P. (1998). Local velocity representation: evidence from motion adaptation. *Vision Research, 38*(24), 3899-3912.

Schwartz, O., Hsu, A., & Dayan, P. (2007). Space and time in visual context. *Nature Reviews Neuroscience, 8*(7), 522-535.

Schwartz, O., Sejnowski, T. J., & Dayan, P. (2009). Perceptual organization in the tilt illusion. *Journal of Vision, 9*(4).

Snowden, R. J., & Milne, A. B. (1997). Phantom motion aftereffects - Evidence of detectors for the analysis of optic flow. *Current Biology, 7*(10), 717-722.

Snowden, R. J., Treue, S., Erickson, R. G., & Andersen, R. A. (1991). The Response of Area Mt and V1 Neurons to Transparent Motion. *Journal of Neuroscience, 11*(9), 2768-2785.

Snowden, R. J., & Verstraten, F. A. (1999). Motion transparency: making models of motion perception transparent. *Trends Cogn Sci, 3*(10), 369-377.

Solomon, S. G., Peirce, J. W., Dhruv, N. T., & Lennie, P. (2004). Profound contrast adaptation early in the visual pathway. *Neuron, 42*(1), 155-162.

Stevenson, I. H., Cronin, B., Sur, M., & Kording, K. P. (2010). Sensory Adaptation and Short Term Plasticity as Bayesian Correction for a Changing Brain. *Plos One, 5*(8).

Stocker, A., & Simoncelli, E. (2006). Sensory adaptation within a Bayesian framework for perception. *Advances in neural information processing systems, 18,* 1289.

Stocker, A. A., & Simoncelli, E. P. (2009). Visual motion aftereffects arise from a cascade of two isomorphic adaptation mechanisms. *Journal of Vision, 9*(9), -.

Takeuchi, T. (1998). Effect of contrast on the perception of moving multiple Gabor patterns. *Vision Research, 38*(20), 3069-3082.

Tanaka, K., Fukada, Y., & Saito, H. A. (1989). Underlying Mechanisms of the Response Specificity of Expansion Contraction and Rotation Cells in the Dorsal Part of the Medial

Superior Temporal Area of the Macaque Monkey. *Journal of Neurophysiology, 62*(3), 642-656.

Tanaka, K., & Saito, H. A. (1989). Analysis of Motion of the Visual-Field by Direction, Expansion Contraction, and Rotation Cells Clustered in the Dorsal Part of the Medial Superior Temporal Area of the Macaque Monkey. *Journal of Neurophysiology, 62*(3), 626-641.

Thompson, P., & Burr, D. (2009). Visual aftereffects. *Curr Biol, 19*(1), R11-14.

Tootell, R. B. H., Reppas, J. B., Dale, A. M., Look, R. B., Sereno, M. I., Malach, R., et al. (1995). Visual-Motion Aftereffect in Human Cortical Area Mt Revealed by Functional Magnetic-Resonance-Imaging. *Nature, 375*(6527), 139-141.

Treue, S., Hol, K., & Rauber, H. J. (2000). Seeing multiple directions of motion-physiology and psychophysics. *Nat Neurosci, 3*(3), 270-276.

Troje, N. F., Sadr, J., Geyer, H., & Nakayama, K. (2006). Adaptation aftereffects in the perception of gender from biological motion. *J Vis, 6*(8), 850-857.

van der Smagt, M. J., Verstraten, F. A., & van de Grind, W. A. (1999). A new transparent motion aftereffect. *Nat Neurosci, 2*(7), 595-596.

van Santen, J. P. H., & Sperling, G. (1985). Elaborated Reichardt Detectors. *Journal of the Optical Society of America a-Optics Image Science and Vision, 2*(2), 300-321.

Verstraten, F. A., Fredericksen, R. E., & van de Grind, W. A. (1994). Movement aftereffect of bi-vectorial transparent motion. *Vision Res, 34*(3), 349-358.

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

Verstraten, F. A. J., Verlinde, R., Fredericksen, R. E., & Vandegrind, W. A. (1994). A Transparent Motion Aftereffect Contingent on Binocular Disparity. *Perception, 23*(10), 1181-1188.

Vidnyanszky, Z., Blaser, E., & Papathomas, T. V. (2002). Motion integration during motion aftereffects. *Trends Cogn Sci, 6*(4), 157-161.

Wall, M. B., & Smith, A. T. (2008). The representation of egomotion in the human brain. *Current Biology, 18*(3), 191-194.

Watson, A. B., & Pelli, D. G. (1983). Quest - a Bayesian Adaptive Psychometric Method. *Perception & Psychophysics, 33*(2), 113-120.

Webster, M. A. (2011). Adaptation and visual coding. *Journal of Vision, 11*(5).

Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature, 428*(6982), 557-561.

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nat Neurosci, 5*(6), 598-604.

Weisstein, N., Maguire, W., & Berbaum, K. (1977). Phantom-Motion Aftereffect. *Science, 198*(4320), 955-958.

Wenderoth, P., & Wiese, M. (2008). Retinotopic encoding of the direction aftereffect. *Vision Res, 48*(19), 1949-1954.

Whitney, D., & Bressler, D. W. (2007). Second-order motion without awareness: Passive adaptation to second-order motion produces a motion aftereffect. *Vision Research, 47*(4), 569-579.

Wilson, H. R., Ferrera, V. P., & Yo, C. (1992). A Psychophysically Motivated Model for 2-Dimensional Motion Perception. *Visual Neuroscience, 9*(1), 79-97.

Wilson, H. R., & Wilkinson, F. (1998). Detection of global structure in Glass patterns: implications for form vision. *Vision Research, 38*(19), 2933-2947.

Winawer, J., Huk, A. C., & Boroditsky, L. (2008). A motion aftereffect from still photographs depicting motion. *Psychological Science, 19*(3), 276-283.

Wissig, S. C., & Kohn, A. (2012). The influence of surround suppression on adaptation effects in primary visual cortex. *Journal of Neurophysiology, 107*(12), 3370-3384.

Wu, S., Lu, H., & Yuille, A. L. (2009). Model selection and parameter estimation in motion perception. In D. Koller, D. Schuurmans, Y. Bengio & L. Bottou (Eds.), *Advances in neural information processing systems* (Vol. 21, pp. 1793-1800). Cambridge, MA: MIT Press.

Xu, H., Dayan, P., Lipkin, R. M., & Qian, N. (2008). Adaptation across the cortical hierarchy: low-level curve adaptation affects high-level facial-expression judgments. *J Neurosci, 28*(13), 3374-3383.

Yuille, A. L., & Grzywacz, N. M. (1988). A Computational Theory for the Perception of Coherent Visual-Motion. *Nature, 333*(6168), 71-74.