UC Irvine
UCI Open Access Publishing Fund

Title
A moving-barber-pole illusion.

Permalink
https://escholarship.org/uc/item/0dc0c197

Journal
Journal of vision, 14(5)

ISSN
534-7362

Authors
Sun, Peng
Chubb, Charles
Sperling, George

Publication Date
2014-05-01

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

A moving-barber-pole illusion

Peng Sun Department of Cognitive Sciences, University of California Irvine, Irvine, CA, USA

Charles Chubb Department of Cognitive Sciences, University of California Irvine, Irvine, CA, USA

George Sperling Department of Cognitive Sciences, University of California Irvine, Irvine, CA, USA

In the barber-pole illusion (BPI), a diagonally moving grating is perceived as moving vertically because of the shape of the vertically oriented window through which it is viewed—a strong shape-motion interaction. We introduce a novel stimulus—the moving barber pole—in which a diagonal, drifting sinusoidal carrier is windowed by a raised, vertical, drifting sinusoidal modulator that moves independently of the carrier. In foveal vision, the moving-barber-pole stimulus can be perceived as several active barber poles drifting horizontally but also as other complex dynamic patterns. In peripheral vision, pure vertical motion (the moving-barber-pole illusion [MBPI]) is perceived for a wide range of conditions. In foveal vision, the MBPI is observed, but only when the higher-order modulator motion is masked. Theories to explain the BPI make indiscriminable predictions in a standard barber-pole display. But, in moving-barber-pole stimuli, the motion directions of features (e.g., end stops) of the first-order carrier and of the higher-order modulator are all different from the MBPI. High temporal frequency stimuli viewed peripherally greatly reduce the effectiveness of higher-order motion mechanisms and, ideally, isolate a single mechanism responsible for the MBPI. A three-stage motion-path integration mechanism that (a) computes local motion energies, (b) integrates them for a limited time period along various spatial paths, and (c) selects the path with the greatest motion energy, quantitatively accounts for these high-frequency data. The MBPI model also accounts for the perceived motion-direction in peripherally viewed moving-barber-pole stimuli that do and do not exhibit the MBPI over the entire range of modulator (0–10 Hz) and carrier (2.5–10 Hz) temporal frequencies tested.

Introduction

Form influences motion perception

It is widely accepted that the visual cortex contains two segregated functional streams of processing. One stream concerns the assessment of spatial relationships, such as motion and location processing, and the other concerns attributes leading to object identification, such as color and form processing (DeYoe & Van Essen, 1988; Maunsell & Newsome, 1987; Ungerleider & Mishkin, 1982). Under this view, the processing of motion information should be independent from the processing of form information. From a theoretical perspective, motion can be computed without explicit form constraints (DeYoe & Van Essen, 1988; Marr, 1982). Indeed, there have been successful motion models that do not concern form information at all, yet manage to explain a wide range of motion perception phenomena (Adelson & Bergen, 1985; Lu & Sperling, 1995; Van Santen & Sperling, 1984, 1985; Watson & Ahumada, 1985; Wilson, Ferrera, & Yo, 1992). However, growing evidence now suggests that form information can influence the extraction of motion information in various ways. (Badcock, McKendrick, & Ma-Wyatt, 2003; Burr & Ross, 2002; Edwards & Crane, 2007; Geisler, 1999; Geisler, Albrecht, Crane, & Stern, 2001; Kourtzi & Kanwisher, 2000; Krekelberg, Dannenberg, Hoffmann, Bremmer, & Ross, 2003; Mather, Pavan, Bellacosa, & Casco, 2012; Pavan et al., 2011; Pavan, Marotti, & Mather, 2013; Ross, Badcock, & Hayes, 2000). The influence of form on motion perception might occur at several different motion processing stages, including at a local motion sensing stage such as V1 (Geisler et al., 2001), at a motion integration stage such as MT (Mather et al., 2012), and even at higher computational levels (Pavan et al., 2013; see Mather, Pavan, Marotti, Campana, & Casco, 2013, for a review).

The barber-pole illusion

Although computational theories of the influence of form on motion perception have been formulated only

Citation: Sun, P., Chubb, C., Sperlin, G. (2014). A moving-barber-pole illusion. *Journal of Vision*, *14*(5):1, 1–27, http://www.journalofvision.org/content/14/5/1, doi:10.1167/14.5.1.

doi: 10.1167/14.5.1 Received October 31, 2013; published May 1, 2014 ISSN 1534-7362 © 2014 ARVO

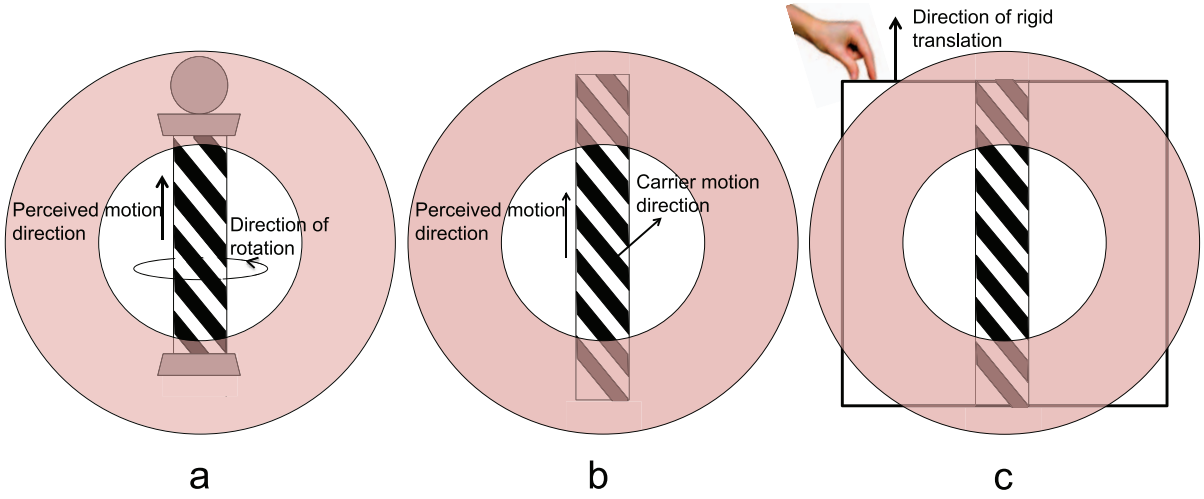

Figure 1. Three ways of producing classical barber-pole displays viewed through a circular window. (The red annulus surrounding the window is actually opaque and contiguous with the background; it is illustrated as being partially transparent to reveal how the displays are created.) (a) View a real barber pole behind a circular window. The cylinder on which the stripes are painted rotates around a vertical axis. Although the real motion in the image plane is purely horizontal, the stripes appear to move vertically upwards. (b) Either on a computer monitor or with real materials, produce a rectangular aperture (the modulator) with the long side oriented vertically. Behind the rectangular aperture, move a grating (the carrier) diagonally upward to the right. Although the real motion is upward to the right, the direction of apparent motion within the aperture is upward, as in (a). (c) On a piece of white paper, produce an image of a barber pole as illustrated, i.e., a snapshot of a barber pole. Drag the snapshot vertically upwards. When a dynamic stimulus can be produced simply by moving a (nondeforming) snapshot of the stimulus, it is called a rigid translation.

quite recently (e.g., Grossberg, Mingolla, & Viswanathan, 2001), the phenomena caused by form-motion interaction were noted long ago. Perhaps the earliest and best-known phenomenon is the barber-pole illusion (BPI; Wallach, 1935), in which a cylinder is painted with diagonal black and white stripes. When the cylinder rotates about its axis, the stripes appear to move parallel to the axis (Figure 1a). In experimental settings, the barber-pole display is usually produced by multiplying a moving grating (the carrier) by a rectangular modulator, so that the motion stimulus is enclosed within the rectangular region (Figure 1b). Like the illusion produced by a real barber pole, the apparent motion direction of the translating bars is parallel to the long axis of the modulator. The illusion cannot be explained by any motion models that do not explicitly consider the shape of the spatial structure enclosing the motion signal. To account for the form-motion interaction that is implicit in the BPI would require further elaboration of the existing motion theories. We now consider previously proposed theories of form-motion interaction.

End-stop theories of the BPI

Obviously, different theories on the BPI can lead to different implications of the underlying neural mechanisms and different implications of the visual processing stages where the form-motion interaction takes place. Perhaps the current predominant view is that the BPI is determined by the motion of bar terminators (or end-stops) at the modulator boundary (Castet, Charton, & Dufour, 1999; Fisher & Zanker, 2001; Kooi, 1993; Lidén & Mingolla, 1998; Lorenceau et al., 1993; Shimojo, Silverman, & Nakayama, 1989). In the end-stop theories, the detection of the end-stop motion is implemented by a designated mechanism often referred as the end-stop mechanism. Motion produced by the end-stop mechanism then combines with that computed by the conventional motion mechanisms to produce the BPI (Tsui, Hunter, Born, & Pack, 2010). The end-stop theory explains the BPI within a purely motion (dorsal) pathway (Pack, Gartland, & Born, 2004; Pack, Livingstone, Duffy, & Born, 2003) and does not require subsequent crosstalk between form and motion processing streams (Mather et al., 2013).

Motion-streak theories of the BPI

Contrary to end-stop explanations, the BPI retains when end-stop motions are in directions different from the modulator's elongated orientation (Beutter, Mulligan, & Stone, 1996; Castet & Wuerger, 1997). Also, the BPI is weakened substantially when the overall end-stop motions are still along the modulator's elongated orientation but the modulator boundary contains irregular details (Badcock et al., 2003). Based on these results, Badcock et al. (2003) proposed that the modulator boundary acted like the motion streak in motion-streak models (e.g., Geisler, 1999) to produce the

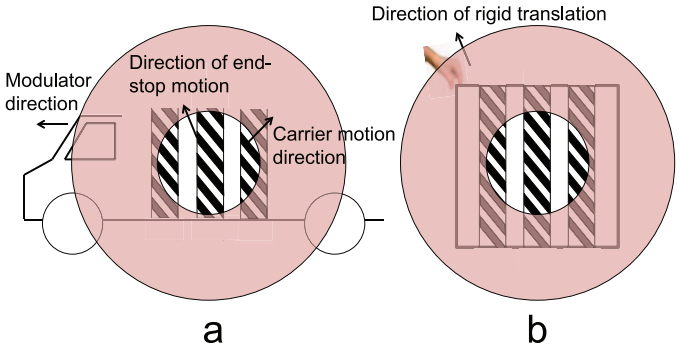

Figure 2. Cartoon illustration of the moving-barber-pole display. (a) A realistic scenario in which a moving truck carries three active barber poles. The truck and the carrier gratings inside each pole move independently. In this cartoon illustration, the truck moves horizontally to the left (the modulator) and the gratings within the barber poles (the carrier) translate diagonally up to the right. Therefore, motion of the bar terminators or end-stop motion is up to the left. Although the stimuli in the experiments are generated as the product of a modulator sinewave grating and a carrier sinewave grating viewed through a Gaussian window, they appear very much as the barber poles on moving truck would appear. (b) Remarkably, a stimulus identical to (a) can be produced by moving a snapshot of the barber poles in a particular direction, the direction of rigid translation, which depends on the speed and direction of the modulator and of the carrier.

BPI. Badcock et al.'s proposal implied late interactions between the form and motion processing streams. Currently, this theory is heuristic, not computational.

Feature-tracking theories of the BPI

A relatively easy but often implicitly articulated theory of the BPI is a feature-tracking theory. Marshall (1990) explained the BPI in a way that is equivalent to tracking the two-dimensional (2-D) spatial features—the bar segments—in Figure 1b. Computationally, the feature-tracking theory is similar to the end-stop theory insofar as end-stops are considered as features.

BPI: An implicit computation of the direction of rigid translation?

The computation of rigid direction is another possible candidate theory for the BPI. We define "rigid direction" as follows. Consider a visual stimulus that is painted or photographed on a sheet of paper; the sheet of paper is translated in a particular direction, then viewed through a window. The direction in which the paper moves is the rigid direction of the visual stimulus. A dynamic display has a rigid direction if and only if all its features move in precisely the same direction at precisely the same speed. The common direction of all the features is the rigid direction of the dynamic pattern. Obviously, only a tiny subset of dynamic patterns has a unique rigid direction. A barber-pole stimulus like the one shown in Figure 1b viewed through a circular window produces exactly the same dynamic stimulus as is produced by dragging a snapshot of the same stimulus upward behind the circular window (Figure 1c). That is, the rigid direction of a barber-pole stimulus is the same as the modulator's elongated orientation, as long as the modulator extends beyond the circular aperture.

A number of different algorithms have been proposed to extract the rigid direction of a moving image. Therefore, motion models with the components that implement these algorithms (Adelson & Movshon, 1982; Heeger, 1987; Perrone, 2004, 2012; Simoncelli & Heeger, 1998) can potentially account for the BPI. These models do not require an end-stop processing mechanism or a feature-tracking mechanism.

The moving-barber-pole display

Is the BPI a result of computing a rigid-motion direction? Or a computation involving tracking features such as end-stops? Or a form-motion interaction that requires late crosstalk between the motion and form processing pathways? It is difficult to discriminate between these theories because for a typical barber-pole display, the feature direction, the rigid direction, and the modulator's elongated orientation are the same, and all are consistent with the direction of the BPI.

A cartoon illustration

Because the different theories of the BPI all make identical motion predictions in a simple barber-pole display, we propose a new diagnostic display: moving barber poles. A typical barber-pole display contains a static modulator window. In a moving-barber-pole display, the modulator moves in a direction and with a speed that is independent of the carrier motion. To illustrate this new display, consider three active barber poles being carried by a truck that is driving leftward while the barber poles it is carrying are just as active as they normally are (Figure 2a). Figure 2b illustrates a counter-intuitive fact: When viewed through a circular window, a view identical to the view of the barber poles being carried by the truck could have been produced by translating a snapshot of the barber poles in a particular direction (i.e., the direction of rigid translation) that depends on the relative speed of the truck and of the barber-pole stripes.

Constructing the moving-barber-pole stimulus

Figure 2 illustrates the basis of the moving-barber-pole stimuli. In the experiments, the moving-barber-

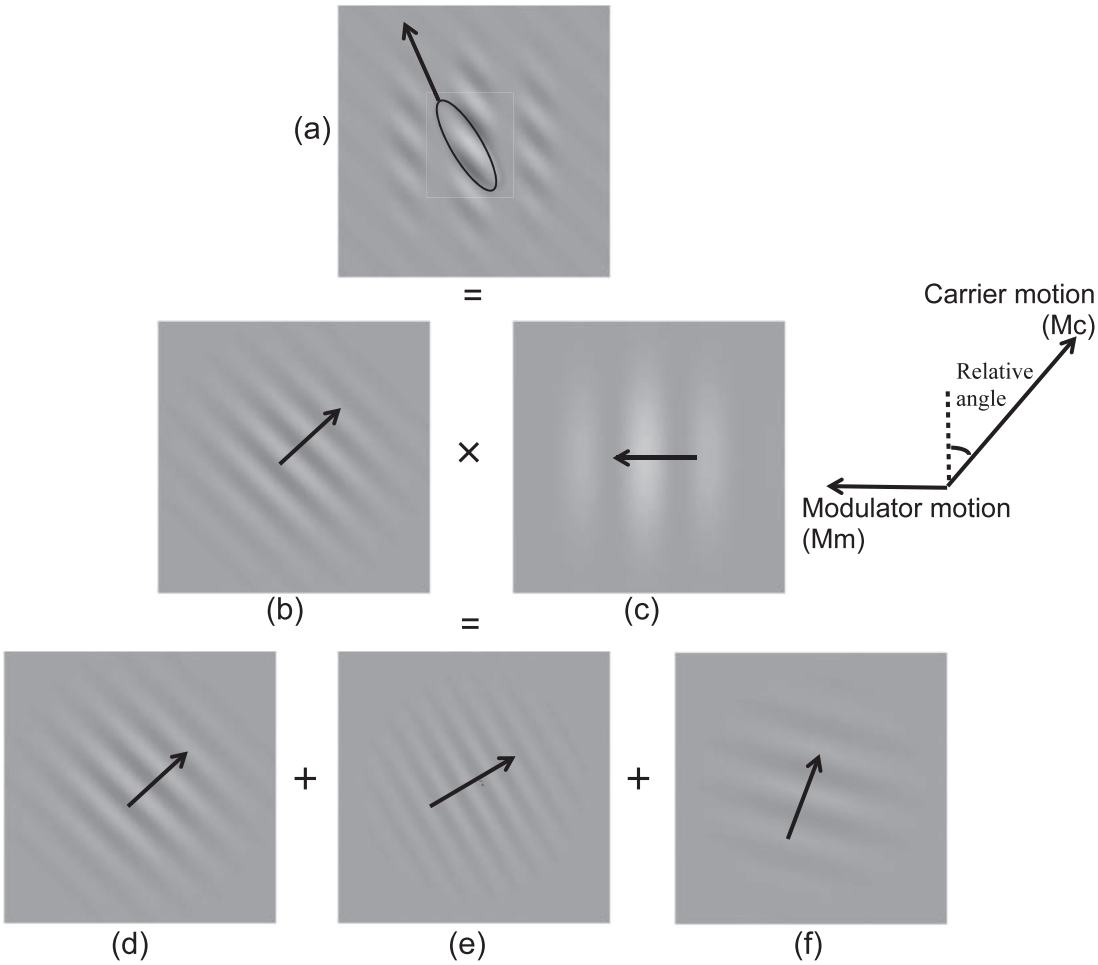

Figure 3. Stimulus decomposition. (a) Snapshot of a moving-barber-pole stimulus. The direction of motion of 2-D spatial features such as end-stops is, by definition, the same direction as the rigid motion of the image within the circular (Gaussian) window. This particular stimulus was constructed by multiplying (b) a moving sinusoid carrier with (c) a moving raised cosine modulator. The multiplication is equivalent to the addition of (d) the carrier grating (same as [b]) and the two side-band components (e). (f). The length of the arrows is arbitrary; arrows merely indicate directions of motion.

pole display is generated by multiplying a moving sinusoidal grating (carrier) with a raised cosine function (modulator) that moves independently of the carrier. Figure 3a shows a snapshot of a moving-barber-pole stimulus. In this example, the carrier moves diagonally up to the right and the modulator moves horizontally to the left. The rigid direction is not in line with the modulator's elongated orientation. By definition, features move in the direction of rigid translation. Therefore, the direction of the feature motion (i.e., the motion of a bar segment as a whole, or end-stop motion) is also different from the modulator's elongated orientation. That the elongation motion direction and the rigid motion direction are quite different makes the moving-barber-pole display highly diagnostic.

Figure 3b through f illustrate the construction of the moving-barber-pole stimulus. Let ω_c and ω_m be the spatial frequencies of the carrier and modulator respectively, let ω_{tc} and ω_{tm} be the temporal frequencies of the carrier and modulator respectively, and let θ_c and θ_m be the angles of the two gratings relative to the vertical, upward direction. Then a moving-barber-pole stimulus is generated by the following equation:

$$\begin{aligned} S(x,y,t) &= \sin(\omega_c(x\cos\theta_c - y\sin\theta_c) + \omega_{tc}t)(1 \\ &\quad + \cos(\omega_m(x\cos\theta_m - y\sin\theta_m) + \omega_{tm}t)) \\ &= \sin(\omega_c(x\cos\theta_c - y\sin\theta_c) + \omega_{tc}t) \\ &\quad + \frac{1}{2}\sin(x(\omega_c\cos\theta_c + \omega_m\cos\theta_m) \\ &\quad - y(\omega_c\sin\theta_c + \omega_m\sin\theta_m) \\ &\quad + (\omega_{tc} + \omega_{tm})t) \\ &\quad + \frac{1}{2}\sin(x(\omega_c\cos\theta_c - \omega_m\cos\theta_m) \\ &\quad - y(\omega_c\sin\theta_c - \omega_m\sin\theta_m) \\ &\quad + (\omega_{tc} - \omega_{tm})t) \end{aligned} \tag{1}$$

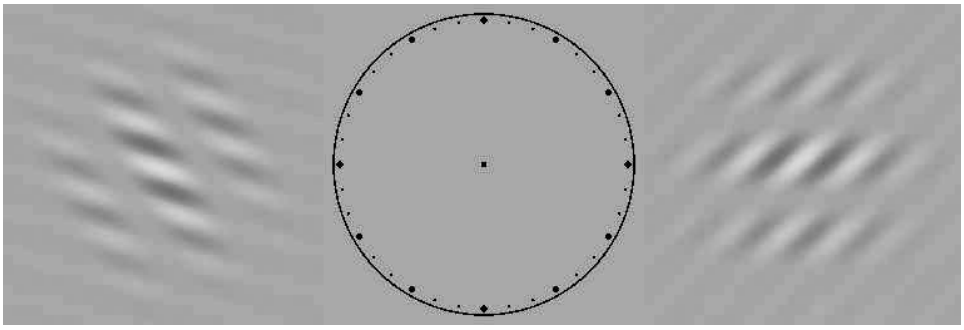

Figure 4. Illustration of the peripheral viewing condition. Within a trial, subjects fixated at the central dot while the stimulus was played for 500 ms. The black clock face remained present during the entire experiment to enforce fixation. Two stimuli are shown here, but only one of these two randomly appeared in any particular trial. In the experiments the orientation of the modulator was randomly chosen between 0° and 359°.

Equation 1 shows that the stimulus can be decomposed into three Fourier components: the original carrier component and two side-band components at half of the carrier contrast. Knowing the Fourier components allows one to examine the role of the first-order motion system, which is often assumed to compute a weighted sum of all available Fourier components.

Introducing the modulator motion adds complications. In particular, the movement of the modulator itself is a higher-order motion signal. In an attempt to isolate a single or smaller number of mechanisms, the observations in this study were primarily peripheral, because peripheral viewing involves fewer higher-order motion computations (Chubb & Sperling, 1989; Lu & Sperling, 1999).

We used multiple barber poles (versus a single barber pole) so that the center of gravity of all the visible material in the viewing window would remain approximately constant and independent of temporal and spatial frequencies of the carrier and modulator. Preliminary observations indicated that some observers were sensitive to the overall movement of a single barber pole. The lateral movement of a single barber pole apparently engages a motion mechanism different from the mechanism that detects movement of the barber-pole stripes (the carrier). The multiple barber-pole stimulus represents our attempt to reduce the number of motion mechanisms involved and thereby, hopefully, to isolate a single motion mechanism.

General methods, all experiments

Apparatus

The experiment was controlled by a Macintosh Intel computer running Matlab (MathWorks, Natick, MA) with the Psychtoolbox package (Brainard, 1997). Stimuli were displayed on a 15-in. Mitsubishi Diamand Pro 710S VGA monitor (Mitsubishi Electric, Tokyo, Japan) with 1024 × 768 resolution running at 85 Hz refresh rate. A lookup table containing 256 gray levels was generated by a standard calibration procedure. The mean luminance of all the stimuli was set at 76.7 cd/m^2.

Subjects

Three naive subjects and one author (S2), ages 22–29, participated in the experiments. Two of the naive subjects (S1 and S4) were psychology undergraduates unassociated with the lab and were kept naive to the purpose of the experiment. A student from a different department was the other naive subject (S3). All methods were approved by the University of California Irvine Institutional Review Board.

Procedure

Peripheral viewing

In Experiments 1, 2, 3, and 4, stimuli were viewed 7° peripherally. Subjects initiated the experiment by pressing a button and were subsequently shown a clock face in the center of the screen (Figure 4). After a fixation period of 1 s, a stimulus could appear on either side of the clock face, the center of the stimulus was always 7° to the left or right of the fixation point. The clock remained on the screen during the whole experiment to enforce fixation. Following an exposure time of 500 ms, subjects indicated the perceived direction of dominant motion on the clock face with a mouse-controlled pointer. Before the formal experiments, naive subjects familiarized themselves with the task by conducting 50 practice trials.

Foveal viewing

In Experiment 5, stimuli were viewed foveally. The experimental setting was the same as in the peripheral

viewing condition except that the stimuli appeared in the center of the monitor.

Experiment 1: Baseline peripheral conditions

The first experiment aimed to establish the baseline performance for peripherally viewed moving barber poles with different modulator temporal frequencies. The orientation of the aperture window remained constant. The carrier temporal frequency was fixed at 10 Hz and the modulator varied. As it did, so did the feature and rigid motion directions. Insofar as the perceived motion directions were in line with the modulator orientation (and remained so irrespective of the variation of the rigid direction with modulator frequency), then this form-motion interaction could not be entirely due to the perceptual computation of the rigid (or feature) motion directions.

Stimuli

Stimuli contained a 10 Hz moving sinusoidal grating (carrier) whose contrast was modulated by a raised cosine function (modulator) that was either static or moving at variable temporal frequencies. When the angle between the carrier's and modulator's motion directions was greater than $\pi/2$, we say that they moved in opposite directions. This configuration resulted in multiple stripes of moving barber poles. The carrier and modulator had a spatial frequency of 1.0 and 0.5 c/d respectively. A Gaussian window with a standard deviation of 1.4° of visual angle was imposed, making the visible area subtend 5° or so. The highest contrast in the stimuli was fixed to 0.4°. The direction of the carrier motion and the orientation of the modulator gratings form an angle that we term the relative angle. For instance the snapshot used in Figure 3 has a relative angle of $-\pi/4$. We tested three relative angles ($-\pi/5$, $-\pi/4$, $-3\pi/10$) and seven different modulator temporal frequencies (−10, −5, −2.5, 0, 2.5, 5, 10 Hz) with the negative sign representing the condition in which the modulator and carrier were moving in opposite directions. For any combination of the factors above, we generated 30 repetitions in which the entire display was randomly rotated. The rotation angles were drawn without replacement from a set of 30 angles that were jittered around 30 evenly spaced angles around the entire clock. Thus in total there were 630 trials in a full session and 30 measurements for each testing condition.

Results and discussion

For each rotated display, we subtracted from each recorded perceived motion direction (PMD) the corresponding rotation angles so that the resulting PMDs were made relative to the direction of a perfect barber-pole motion, i.e.., relative to the upward vertical direction. Figure 5a shows the histograms of a naive subject's (S1) PMD for a stimulus with relative angle of $-\pi/5$ over all modulator temporal frequencies. Each panel corresponds to a particular modulator temporal frequency. The barber-pole motion direction is at 0° (the upward vertical direction) in all panels, whereas the rigid direction changes as the modulator temporal frequency changes. Circular Gaussian functions (Berens, 2009) were then fit to the PMDs for each modulator temporal frequency, and the mean values of those circular Gaussian functions were plotted as functions of temporal frequencies, as shown in Figure 5b. Directions of the carrier motion (solid blue), rigid motion (solid green), higher-order motion (dotted red), and barber-pole motion (dashed red at 0°) are annotated to be compared against the mean PMDs (black). In the same format, we present the data for all subjects under all testing conditions in Figure 6.

The term MBPI is used to refer to the phenomenon in which the perceived motion direction of a moving-barber-pole display is inline with the modulator's orientation. For a range of modulator temporal frequencies (variable across subjects), the PMD curves in Figure 6 overlap the moving-barber-pole curve, i.e., they exhibit the MBPI. One naive subject (S4) showed relatively small ranges of MBPI. For all tested temporal frequencies, PMDs always fell in between the rigid direction and the MBPI direction. For most subjects, PMDs coincided with the rigid direction only at high temporal frequencies of the modulator. The clear deviation from rigidity shows that the MBPI requires a computation other than tracking features such as end-stops or computing the rigid direction.

First-order carrier motion alone does not explain the result. However, the carrier motion is not the only Fourier component available to the first-order system (Figure 3d through f), although it should be the dominant one due to the fact that its contrast is twice the contrasts of the two side-band components.

Sperling and Liu (2009) proposed a quantitative theory of the first-order motion system processing that applied to Type 1 and Type 2 (Wilson et al., 1992), plaid stimuli with equal spatial frequencies. When only the first-order system was involved, perceived motion directions of their foveally viewed plaid stimuli were accurately predicted by a linear summation of the two components' first-order motion-strength vectors.

The magnitudes of the two vectors were proportional to the squared values of the component's contrasts. The

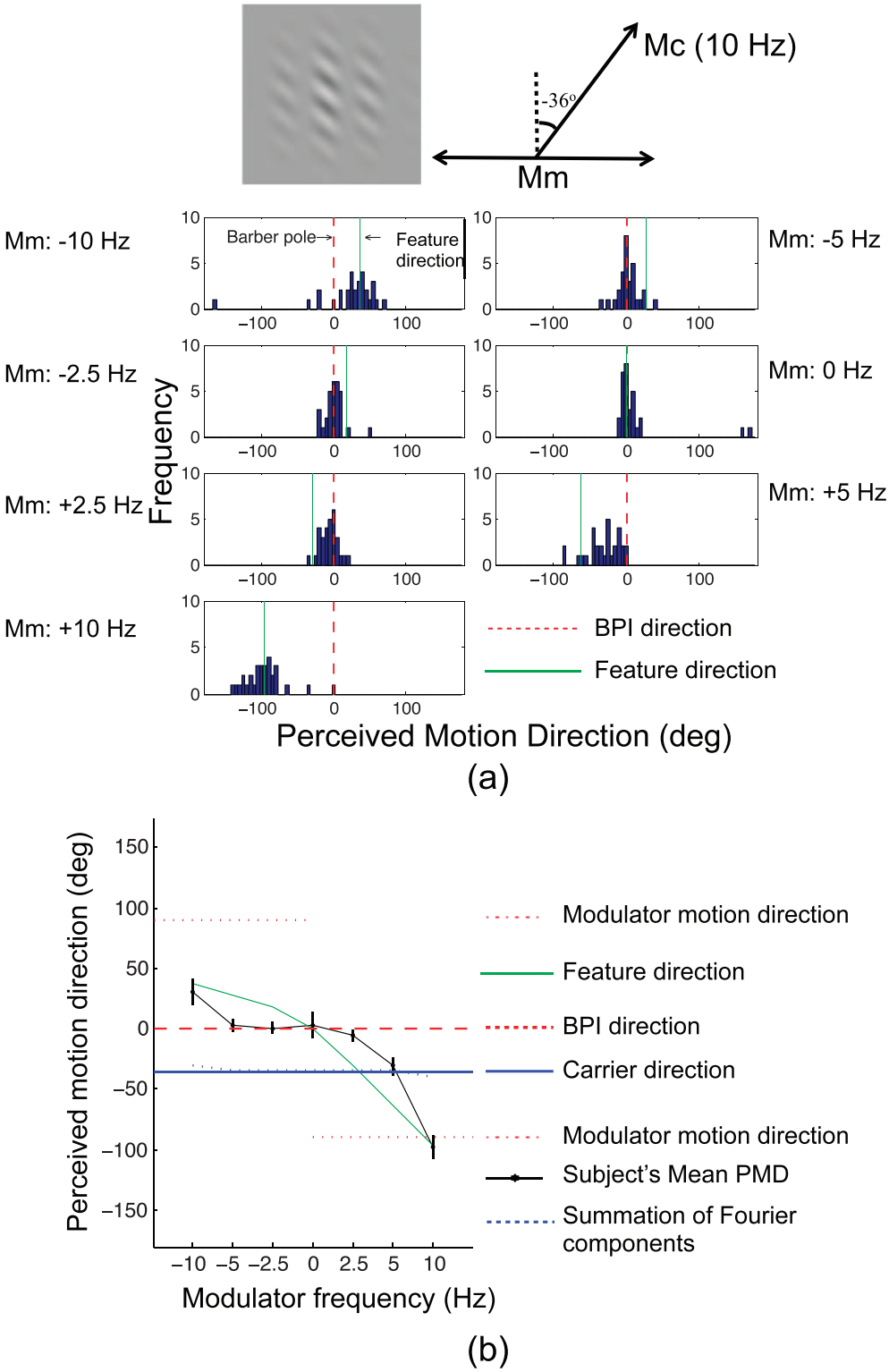

Figure 5. Experiment 1: An example of a naive subject's perceived direction in moving-barber-pole stimuli with modulator temporal frequencies ranging from −10 to +10 Hz. (a) The histograms of PMD for a moving-barber-pole stimulus that has a relative angle of –/5, versus modulator temporal frequency. The green solid line and the dashed red line, respectively, indicate the feature direction and the MBPI motion direction. Although on each trial the rotation angle was chosen randomly between 0° and 359° deg, here the MBPI motion direction is always shown as 0°, corresponding to the upward vertical direction in the stimulus snapshot. (b) The mean value of each PMD histogram plotted against modulator temporal frequency; the carrier temporal frequency was 10 Hz. Error bars represent 95% confidence intervals.

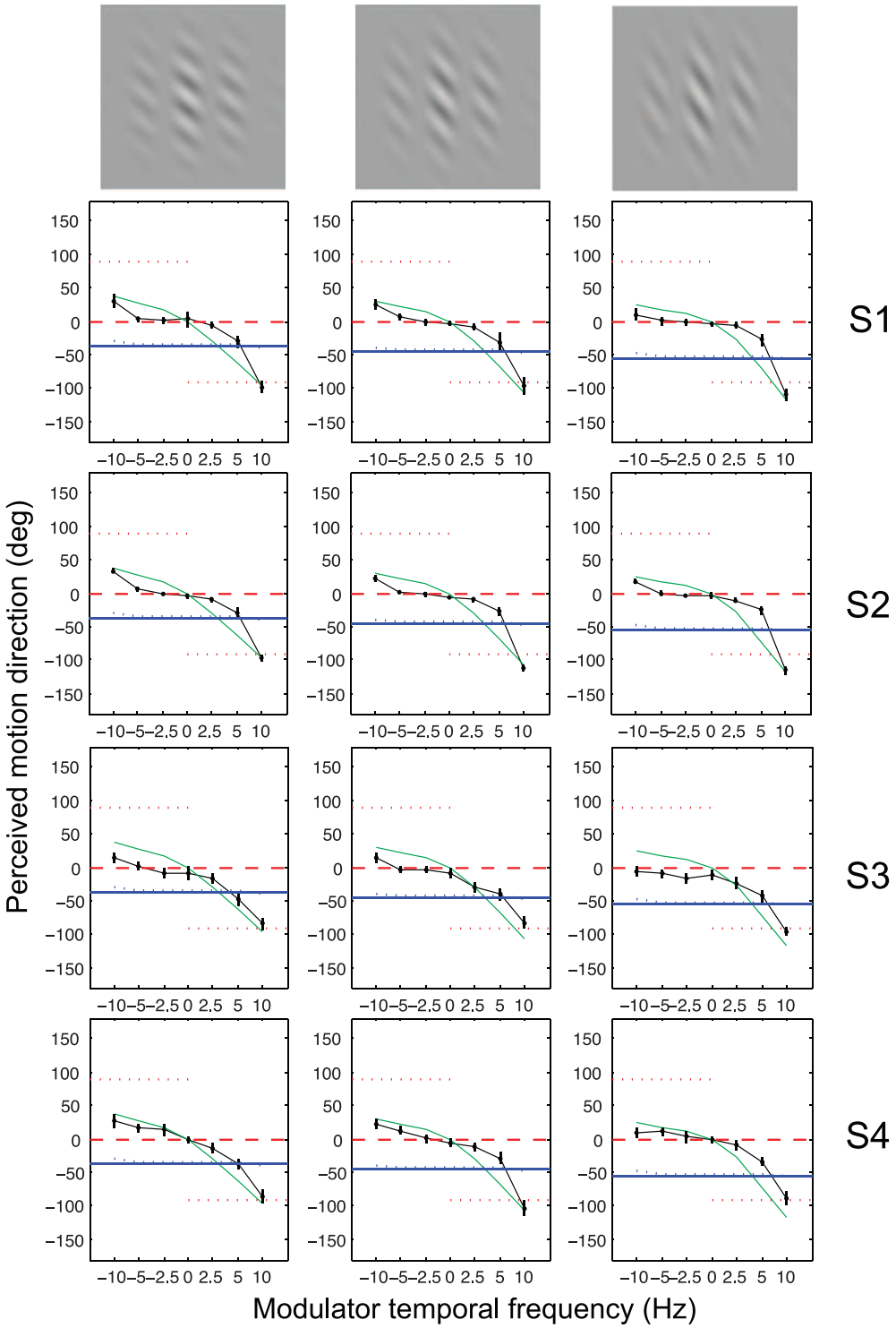

Figure 6. Experiment 1: Mean perceived motion directions of four subjects (separated in rows) in peripherally viewed moving-barber-pole stimuli of three relative angles (separated in columns). Notation is the same as Figure 5b.

predicted motion direction of this theory is illustrated by the dotted blue line in Figure 5b, which is very close to the carrier motion direction and therefore does not explain the results.

One possible explanation for the MBPI is that the visual system conducts a vector summation of the carrier (first-order) and modulator (higher-order) motions. It is possible that the strength of the modulator motion is reduced in the periphery such that the remaining higher-order motion strength and the intact strength of the first-order motion add up, coincidentally, to align with the direction of MBPI.

Such a summation of first- and higher-order motion vectors has been proposed in other contexts (Tse & Hsieh, 2006; Wilson et al., 1992).

However, even the data in Experiment 1 alone indicate that the combination of first- and higher-order motion is unlikely to explain MBPI. This is because the MBPI prevailed when the modulator was static. PMD would have been close to pure first-order carrier motion (not MBPI) had it resulted from the combination of first- and higher-order motions vectors. Also, when the carrier and modulator moved in the same direction (i.e., when the modulator temporal frequencies were positive in Figure 5b), PMDs did not fall between the carrier and modulator motion direction as the combination rule would have predicted. Instead, PMDs fell between the MBPI and the feature direction. The MBPI requires another explanation.

Experiment 2: Carrier temporal frequency

Experiment 1 tested the dependency of the MBPI on the higher-order modulator temporal frequency, and found that the MBPI obtained over a middle range of modulator frequencies.

Experiment 2 investigated the effect of varying the carrier temporal frequency.

Stimuli

Stimuli remained the same except that the relative angle of the carrier stripes was fixed at $-\pi/4$ and three more carrier temporal frequencies (5, 2.5, 0 Hz) were included. That is, the previous condition of 10 Hz carrier temporal frequency was interleaved with three more conditions of the carrier temporal frequency. Of course, when the carrier temporal frequency was 0 Hz, the modulator was never 0 Hz.

Results and discussion

Figure 7 shows mean PMDs as functions of modulator temporal frequencies at four carrier temporal frequencies. At the carrier temporal frequency of 10 Hz, PMDs are essentially identical to those in Experiment 1. As carrier temporal frequency decreases, so does the influence of the MBPI direction on the PMD—an indication that MBPI depends on the carrier temporal frequency. When carrier temporal frequency decreases to zero, PMD is entirely in the feature direction. In fact, PMDs always lie between the feature direction and the MBPI motion direction.

As in Experiment 1, increasing modulator temporal frequency tends to align PMDs with the feature direction. Inspection of Figure 7 shows that, although carrier motion and modulator motion both influence PMD, it cannot by simple addition of velocity vectors.

Experiment 3: Phase of adjacent barber poles

Experiment 2 showed that the MBPI was affected by changes in the carrier temporal frequency. The carrier motion in the three visible barber poles within a circular Gaussian window in the stimuli of Experiments 1 and 2 was derived from the same fundamental sine-wave component (Figure 3d). This configuration may generate an implicit computation (or even a percept) of a single, extended sine wave (rather than of several sine waves restricted within each pole) moving behind vertical occluders (the zero-contrast regions), especially in the periphery. Experiment 3 investigated whether having a single carrier sine-wave component fill all three barber poles is critical for the MBPI.

Stimuli

The single sine-wave component was effectively removed by reversing the carrier contrast in half of all the barber poles (see the example in Figure 8). To do this, the sign of the raised cosine modulator was flipped in every other spatial cycle. Within each pole, however, the carrier motion was perfectly retained despite this manipulation.

Results and discussion

PMDs for the control condition and the contrast-reversed condition are shown in Figure 8. The MBPI is dominant in the control condition (left column) and is hardly affected even when the percept of a single carrier motion is not available (right column). This suggests that the computation of the MBPI does not need the computation of the single, extended carrier motion. In fact the whole pattern of the PMD curve seems unaffected. If the computation giving rise to the PMD curve does involve the computation of the carrier motion, then the carrier motion is probably computed locally within each pole, rather than over the entire visible area.

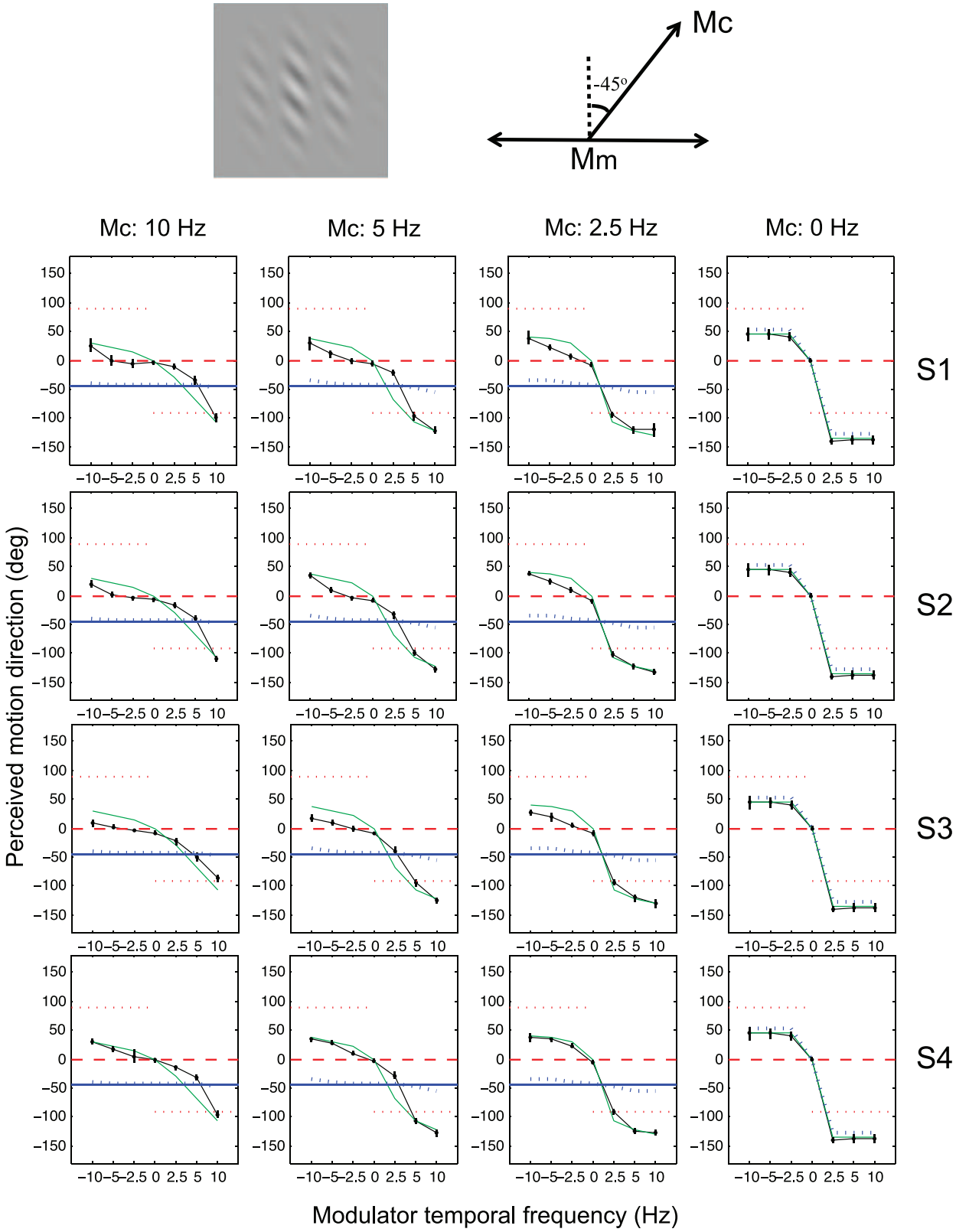

Figure 7. Experiment 2: Mean PMDs of four subjects in peripherally viewed moving-barber-pole stimuli in which the temporal frequency of the carrier was varied from 10 Hz to 0 Hz. Notation is the same as Figure 5b. The condition in which both carrier and modulator were static was not tested; the data point in that case is drawn at (0, 0).

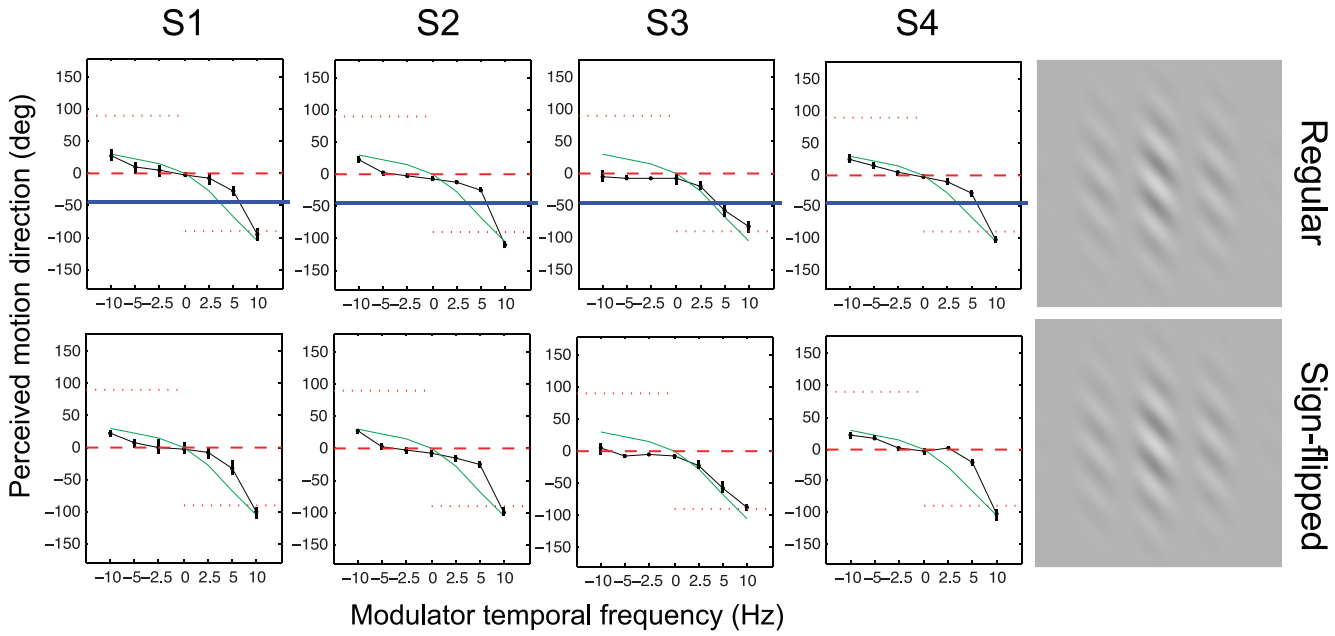

Figure 8. Experiment 3: Same or opposite phase of the carrier in adjacent barber poles. Notation is the same as Figure 5b. The upper row corresponds to the interleaved control condition in which all barber poles share the same fundamental sine-wave carrier, a condition that was tested in Experiments 1 and 2. The lower row shows PMDs in the altered moving-barber-pole stimuli in which the raised cosine modulator flipped signs every other spatial cycle. This manipulation causes the carrier in adjacent barber poles to have reversed contrasts, resulting in the elimination of the fundamental carrier sine-wave component from the spectrum. PMDs do not differ significantly for stimuli with and without the fundamental sine-wave component.

Experiment 4: Masking the modulator

Humans are sensitive to motion that is not based on luminance modulation. Such an ability has been attributed to the existence of second- and/or third-order motion systems in the visual system (Lu & Sperling, 1995). The second-order motion system is based on contrast-energy modulation. The third-order motion system operates on features and has a lower temporal resolution than the second-order system. As the experiments here do not distinguish between second- and third-order motion, we refer to them collectively as higher-order motion. The motion of the translating modulator in the moving-barber-pole stimuli can, in principle, be detected by a higher-order motion system. Experiment 4 investigates the extent to which the MBPI depends on the detection of modulator motion by higher-order motion systems. Masking stimuli are introduced that significantly reduce the power of higher-order motion signals in moving-barber-pole stimuli.

Stimuli

The introduction of masking stimuli was greatly facilitated by replacing the raised cosine modulator of Experiments 1, 2, and 3 with a raised square-wave modulator. The zero contrast regions were filled with phase-randomizing sine-wave gratings of the same orientation as that of the carrier. The masking gratings changed phase randomly every two screen refreshes on the 85 Hz monitor. As a result, each masking strip had a broad temporal frequency band whose range was between 0 and 21.5 Hz in the case of static modulators. Two spatial frequencies of masking gratings were tested: 1.0 c/d (same as the carrier) and 2.0 c/d. For the 2.0 c/d masking grating, one extra contrast condition was included in which the masking grating had twice the contrast of the carrier.

The reason for including an additional double-contrast level for the 2 cpd masking grating was that higher spatial frequency textures were found to produce less energy than lower spatial frequency textures in the context of higher-order motion processing (Werkhoven, Sperling, & Chubb, 1993). Sine-wave gratings of twice the spatial frequency of the carrier cannot silence the higher-order system because of the existence of the energy difference. Thus doubling the contrast of the high spatial frequency masking stimuli will bring the masking stimuli close to the carrier in terms of the texture energies on which the higher-order system operates (Werkhoven et al., 1993). Snapshots of the new stimuli are given in Figure 9. Since our interest in this experiment was whether MBPI depended on higher-order motion processing and therefore could be

(a) (b) (c)

Figure 9. Examples of three masking gratings used in Experiment 4. Moving sine waves are modulated by translating square-wave modulators. Where modulator contrast is zero, masking stimuli made of phase-randomizing sine waves are added. (a) The masking gratings have the same spatial frequency as the original sine-wave carrier. (b) The masking gratings' spatial frequencies are twice those of the original sine wave. (c) Same as (b) but the masking gratings have twice the contrast of the original carrier.

abolished or diminished by the lateral masking, we only tested the five modulator temporal frequencies (−5, −2.5, 0, 2.5, 5 Hz) for which the MBPI was most pronounced.

Results and discussion

Figure 10 shows results for all subjects. Overall, the Experiment 4 data obtained with square-wave barber poles and an interposed masking grating are remarkably similar to the Experiment 1 data obtained with sinusoidal barber poles without masking. Insofar as there may be a slight difference in PMD between masked and unmasked stimuli, it occurs only at the maximum modulator temporal frequency (+5 Hz, at the extreme right of the panels in Figure 10) where there is no MBPI. There, in 4 out of 12 panels, masked PMDs deviate slightly more than unmasked PMDs towards the rigid direction. The obvious conclusion is that these masking gratings have no effect whatsoever on the MBPI. Therefore, varying the modulator temporal frequency in Experiments 1 and 2 must have affected the MBPI without affecting higher-order motion computations. In the periphery, higher-order motion perception of the modulator is already so weak that masking it has no further effect.

Experiment 5: Foveal viewing

The previous four experiments tested exclusively peripheral viewing. Higher-order motion perception is much stronger in the fovea than in the periphery, so it is of interest to determine how subjects would respond to the same stimuli that had been viewed peripherally when they were presented in the fovea. Until this point, the three naive subjects had never seen the testing stimuli in their fovea. Experiment 5 measured the PMDs for these same four subjects when they viewed moving-barber-pole stimuli foveally.

In Experiment 4, the addition of masking gratings did not eliminate the MBPI. In foveal viewing (Experiment 5) higher-order motion (both second-order and third-order) is much stronger than in peripheral viewing (Chubb & Sperling, 1989; Lu & Sperling, 1999). To determine the role of higher-order motion in the foveal MBPI, Experiment 5 also measured their PMDs in moving-barber-pole stimuli in which higher order motion has been significantly attenuated by the addition of masking gratings that have the same spatial frequency as the carrier (as in Experiment 4, Figures 9a and 10, column 1).

Stimuli

Stimuli were a standard moving barber pole with a relative angle of $-\pi/4$ and a barber pole with masking gratings of the same spatial frequency as the carrier (Figure 9a). The two conditions were tested in separate blocks. For comparability with the previous experiments, subjects were required to judge the dominant motion direction. Following the standard testing procedure in which subjects only indicated perceived motion directions, in a separate session, subjects viewed the stimuli foveally and only gave verbal descriptions of the motions they saw.

Results and discussion

For standard, masking-free, moving barber poles (except when the modulators move at 5 Hz in the same direction as the carrier), according to their verbal descriptions, subjects typically saw two separate motions, one corresponding to the higher-order modulator motion and the other corresponding to local motion streams inside each pole. In order to give an

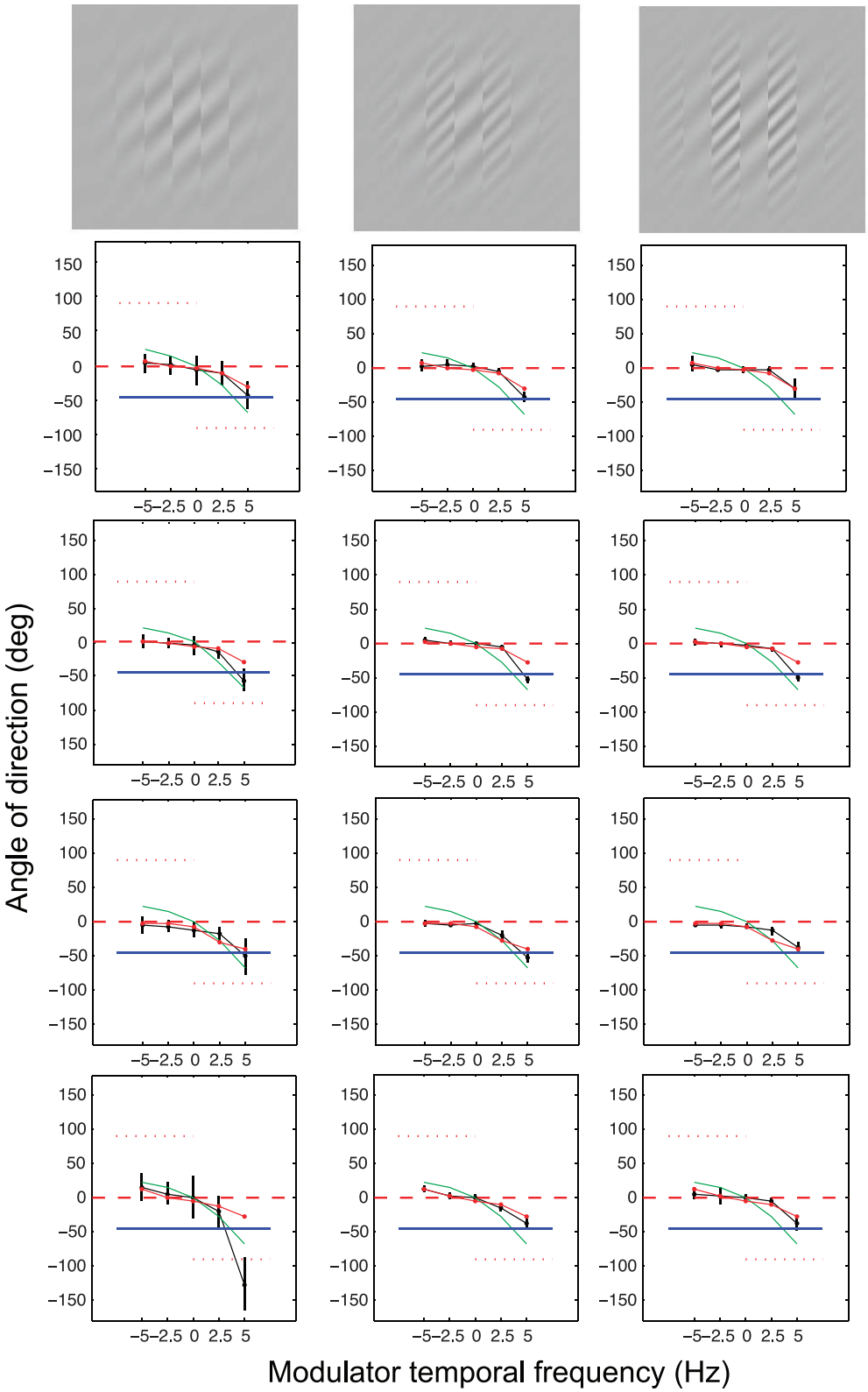

Figure 10. Experiment 4: Mean PMDs of four subjects in peripherally viewed moving-barber-pole stimuli with masking gratings that reduce the influence of higher-order motion-perception mechanisms. Solid black lines connect the data points. The solid red line is comparison data from Experiment 1 (Figure 6, middle column) obtained without masking gratings. Other notation is the same as Figure 5b. Columns are arranged in the same order as in Figure 9.

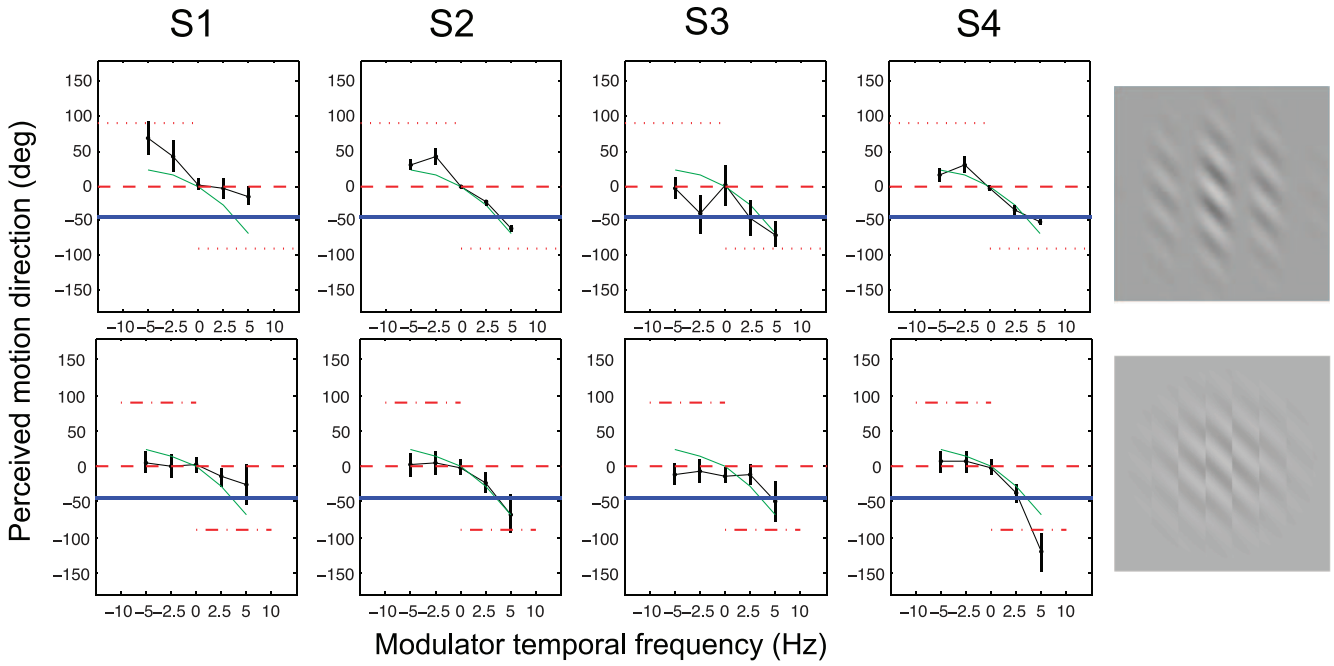

Figure 11. Experiment 5: Mean PMDs for foveally viewed moving barber poles with masking gratings (bottom row) and without (top row).

estimate of only one dominant motion, they variously responded to either one of the motions or they responded with an arbitrary compromise between those motions. When the modulator moved at 5 Hz in the same direction as the carrier, they saw a single coherent motion (the feature, i.e. rigid motion).

PMDs for standard moving barber poles are given in the first row of Figure 11. No obvious MBPI is present compared to the same stimuli viewed peripherally (Figure 6). Interestingly, in some conditions, the mean PMDs are close to the higher-order modulator motion direction, which never occurred in the peripheral viewing condition. Large error bars indicate that PMDs obtained in the fovea are diverse. When plotted as circular histograms (Figure 12), PMDs are indeed very diverse and sometimes show multi-modality patterns. In other words, when the higher-order modulator motion is available to the motion system, it does not help form the MBPI motion but instead masks it. On the other hand, attenuating the higher-order modulator motion with the masking gratings in foveal viewing (Figure 11, bottom row) or by peripheral viewing (Figure 6) greatly facilitates the MBPI.

In contrast to the diverse histograms in foveally viewed standard moving-barber-pole stimuli, with the addition of masking gratings, the circular histograms of PMDs become more compact and centered on the MBPI direction, as shown in Figure 13. This is especially true when the modulator moves in the opposite direction of the carrier. Also notable in Figure 13 is that sometimes PMD is in the reversed direction of MBPI. In some cases this might be due to a confusion of random motion components of the interleaved masking gratings with the motion of the moving-barber-pole grating. However, the frequency of reversed versus normal MBPI directions in several of the panels of Figure 13 suggests the possibility of reversed MBPI.

Note that the mean PMDs for S4 in the lower right panel of Figure 11 are not in the same range as those for the other three subjects. But S4's circular histograms under the two conditions (the rightmost two panels in the bottom row in Figure 13) reveal that this subject frequently perceived a reversed MBPI. The average of MBPI and reverse MBPI gives an unusual angle. Apart from this one case, the other mean PMDs for the masked moving barber poles clearly demonstrate the change of PMDs from an irregular, random pattern (first row) to a MBPI dominated pattern (second row). In other words, the addition of the between pole masking gratings restored MBPI in the foveal view.

Figure 14 shows histograms of PMDs pooled for all subjects. For regular moving-barber-pole stimuli (top row), the component of MBPI is either swamped amidst the broad range of the responses (e.g., for modulator frequency −5 and −2.5 Hz) or is very weak compared to the component of the rigid direction (e.g., for modulator frequency +2.5 and +5 Hz). Of course for 0 Hz, a stationary modulator, i.e., a classical barber pole, a normal BPI is observed. When the modulator is masked (bottom row), PMDs clearly peak around the MBPI direction for almost all modulator temporal frequencies. The exception is +5 Hz modulator

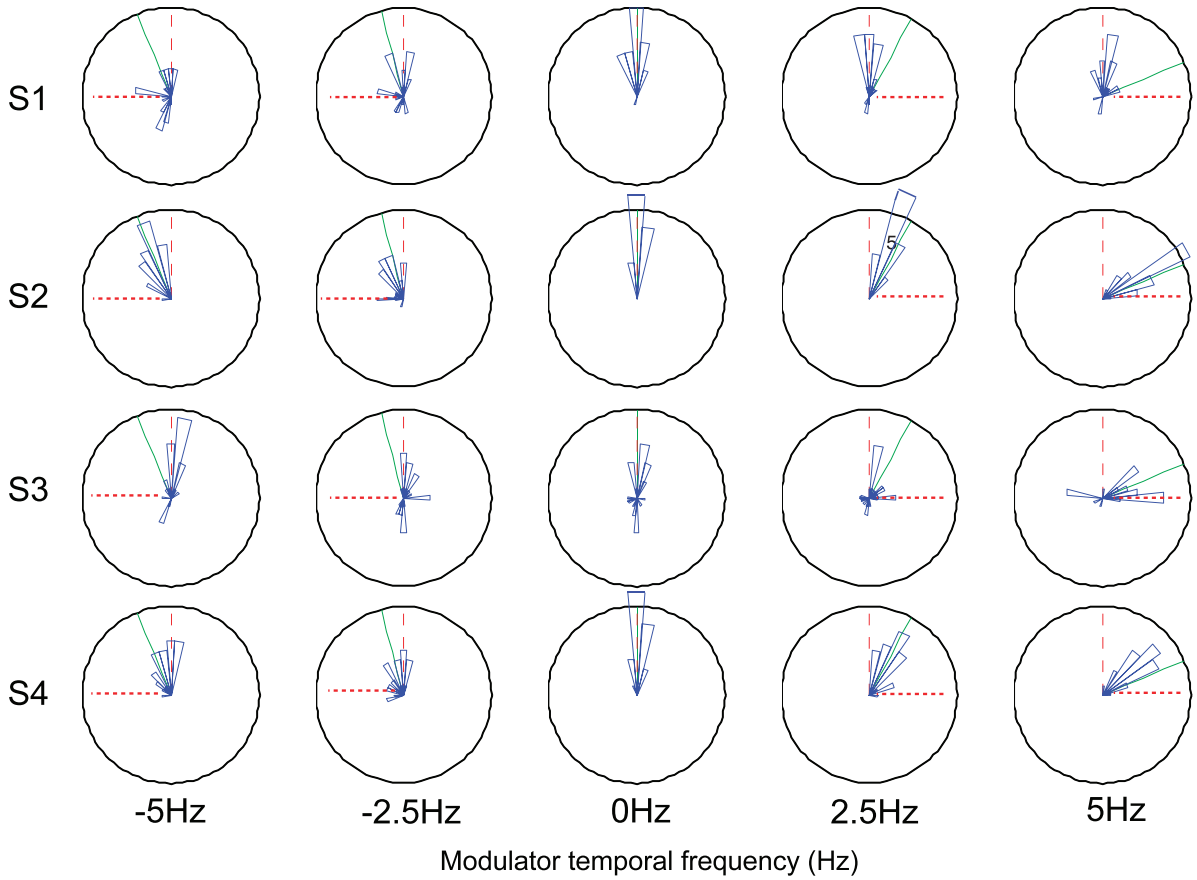

Figure 12. Experiment 5: Circular histograms of PMDs for moving barber poles without masking gratings viewed foveally. Rows represent subjects, columns represent modulator temporal frequencies. Each bin is 10° wide. The boundary of the circle corresponds to 10/30 trials. The dashed red lines represent the MBPI direction. The horizontal dashed red lines represent the modulator direction except at 0 Hz when there is no modulator motion. The solid green lines represent the rigid (feature) direction.

frequency, in which case there is a secondary peak around the reversed MBPI direction.

Even in the case of a stationary modulator, the masking grating compresses the histogram of PMDs, i.e., it improves the classical BPI. This is contrary to the observations of Lalanne (2006) who found that lateral masking and adaptation impaired the BPI. This may be due to the use of a much lower temporal frequency of the carrier (2.5 and 3.3 Hz) than the 10 Hz in the present study. Lower temporal frequencies make possible the involvement of higher-order motion processes in the perception of carrier motion.

Review

In the moving-barber-pole stimulus, the directions of rigid motion (same as feature motion), first-order motion, higher-order motion, MBPI motion are all different. However, the fact that MBPI is consistently perceived over a wide range of parameters is not predicted by any current model or theory. We consider here what current motion perception models do predict as the direction of perceived motion in moving-barber-pole stimuli.

We begin with the most general motion-perception models.

For moving-barber-pole stimuli, models based on the principle of computing the rigid direction (Heeger, 1987; Schrater, Knill, & Simoncelli, 2000) predict the rigid-motion direction, not the MBPI motion direction. We consider the predicted response to moving-barber-pole stimuli of two models of MT pattern cells (Perrone, 2004; Simoncelli & Heeger, 1998). A distinctive feature of MT pattern cells is that they are sensitive to the direction of the corresponding rigid translation of a pattern, rather than directions of Fourier component motions. These MT pattern cell models are sensitive to the rigid direction not the MBPI direction of the moving-barber-pole stimulus. Perrone (2012) used one of the MT pattern cell models (Perrone, 2004) to generate a field of local velocities (speeds and directions), on which further processing could be operated. For the moving-barber-pole stimulus, we were able to show that these velocity vectors all point in

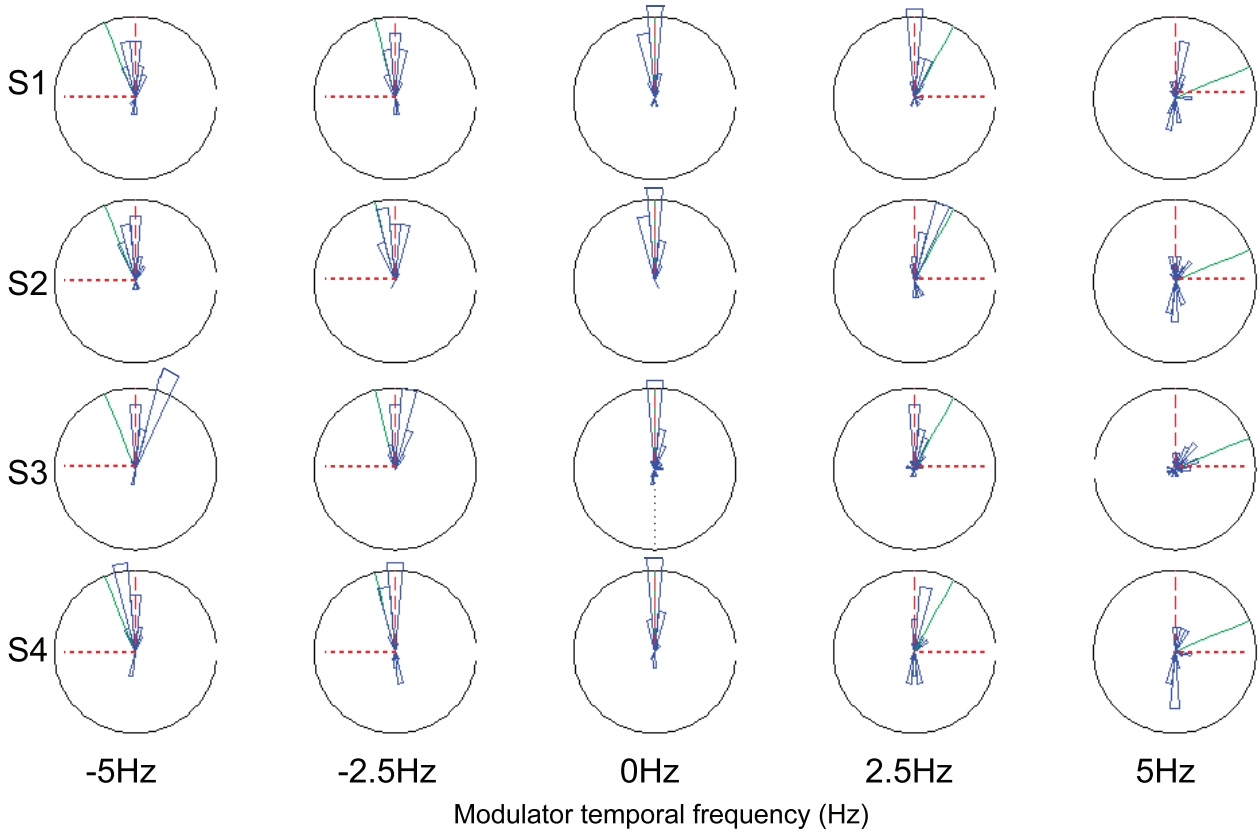

Figure 13. Experiment 5: Circular histograms of PMDs for moving barber poles with masking gratings viewed foveally. Notation is the same as Figure 12.

the rigid direction. The field of MT pattern-cell velocities would require further elaboration to account for the MBPI.

Feature tracking (Bowns, 1996; Rubin & Hochstein, 1993) or end-stop theories (Tsui et al., 2010) also predict rigid motion because, by definition, features move in the rigid direction. Models based on summing first- and higher-order motion vectors (Wilson et al., 1992) are vetoed by our data. This is because in our peripheral stimuli, higher-order motion had no influence on the perceived direction, and the first-order motion direction was not in the MBPI motion direction. Moreover, in foveal viewing, Experiment 5 showed that higher-order motion interfered with the MBPI, and that canceling higher-order motion produced a foveal MBPI. Since none of the general motion perception models works, it seems reasonable to consider mechanisms that specifically deal with the MBPI, i.e., mechanisms that not only compute a motion signal but also take into account the shape of the modulator window or other non-motion aspects of the motion stimulus. We briefly review four models that incorporate such extra mechanisms before proposing a new model.

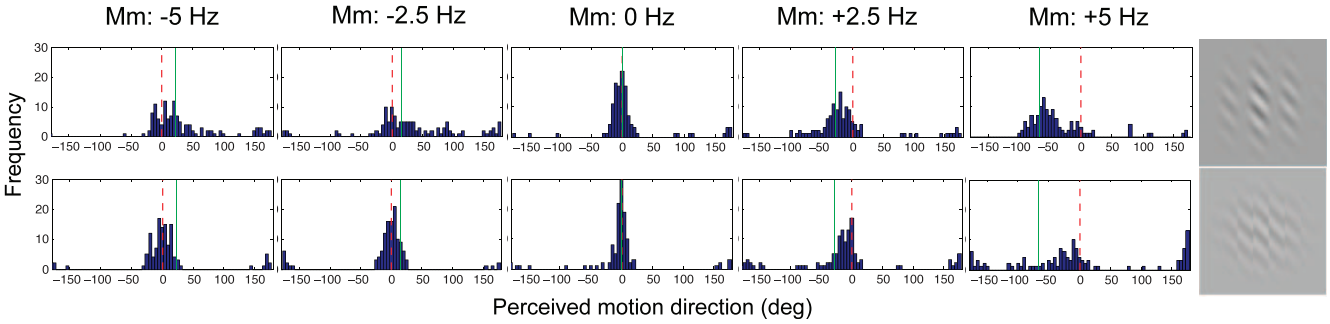

Figure 14. Experiment 5: Histograms of PMDs in foveally viewed moving barber poles for five modulator frequencies, with and without masking, pooled for four subjects. Upper row: without masking, Lower row: with masking.

Noticing that spatial orientation-selective mechanisms were involved in the detection of motion direction, Geisler (1999) proposed a spatial motion-direction sensor that consists of a orientation selective subunit and a motion selective subunit. The subunits were so arranged that the preferred directions of the motion unit and the orientation of the spatial unit were parallel. Outputs of the two subunits were combined multiplicatively to produce a response that peaked when the stimulus motion produced a motion streak oriented parallel to the orientation of the spatial subunit. The anisotropy power spectrum theory proposed by Barlow and Olshausen (2004) extends the spatial orientation subunit in Geisler's model (1999) to include spatial higher-order units. The Barlow-Olshausen theory implies that higher-order, non-Fourier spatial patterns can bias motion perception towards the orientation of higher-order spatial subunit, much as the first-order spatial subunit in Geisler's theory could bias the perceived motion directions towards its spatial orientation. The biasing influence of higher-order spatial patterns was verified by Badcock and Dickinson (2009). This extended spatial motion-direction sensor is consistent with some of our data but it fails to explain why masking the higher-order pattern does not impair (as predicted) the MBPI. Furthermore, it fails to explain why reducing carrier temporal frequency weakened the MBPI.

Grzywacz, Watamaniuk, and Mckee (1995) proposed a temporal coherence theory to explain the enhanced dot motion detection when a target dot moved along a regular trajectory amidst many noise dots undergoing Brownian motion. Central to this theory is the idea that local motion detectors with similar directional tunings are connected along their preferred direction in space, and successively in time—literately a feature-tracking mechanism. Therefore, when the connected detectors in the sequence are all stimulated, as in the case of a dot moving consistently in one direction across several frames, this detector sequence will fire strongly. Essentially a feature tracking model, the temporal coherence theory always predicts motion in the direction of features (i.e., the rigid direction) and therefore cannot be a good theory for our data. Furthermore, in experiments that do not involve isolated dots, the Grzywacz et al. (1995) theory does not specify the relevant features to be matched between frames (the correspondence problem). Nevertheless, Grzywacz et al. (1995) introduces an interesting concept: motion detectors with similar directional tunings connect to one another to respond more strongly to coherently moving objects. Coherent direction tuning is reminiscent of the phenomenon of lateral facilitation found in other sensory domains, such as in spatial contour integration (Field, Hayes, & Hess, 1993), and will be an essential feature of the STI model.

Grossberg et al. (2001) proposed a complicated formotion model to account for various form-motion-interaction phenomenon including the barber-pole illusion. The formotion model cleverly integrates reliable motion signals often located at end-stops and T-junctions with ambiguous motions elsewhere to reproduce the temporal dynamics of the transition from local bar motion to rigid motion as display duration increases (Lorenceau et al., 1993). Of more relevance to our study is its short-range and long-range grouping stage, which echoes the idea that a higher level of motion process combines local motion signals in both space and time as in Van Doorn and Koenderink (1984) and Fredericksen, Verstraten, and Van De Grind (1994a, 1994b). In the formotion model, the long-range grouping is carried out over a Gaussian shaped, isotropic spatial region whereas the short-range grouping is carried out over a small anisotropic region oriented in the preferred motion direction. Therefore, a direct application of the formotion model to our moving-barber-pole stimuli will (incorrectly) predict the rigid motion direction for the moving-barber-pole stimuli regardless of modulator and carrier temporal frequencies.

More recently Bowns (2011) introduced a model that detected collinear patterns in feature traces accumulated over time. This model is designed to explain plaid motion in which two moving sine-wave components are superimposed on top of each other. The model first finds edges by finding zero-crossings of second-derivatives of each individual sine-wave component. Features are then extracted by taking only the intersections (spots) of the two groups of edges. Over time those intersection spots will trace out lines similar to motion streaks. The model then runs a Hough transform to detect dominant lines produced by feature traces and takes the orientations of these dominant lines as the predicted motion direction. Insofar as the model finds features within a frame and traces the motion of these features, it predicts the rigid direction for moving-barber-pole stimuli, not the observed MBPI. Nevertheless, in its Hough transform algorithm, Bowns' model implicitly implements spatial-temporal information integration that is critical for explaining the phenomena observed in moving-barber-pole stimuli.

An intuitive illustration of the motion-path-integration model

We first briefly overview the motion-path-integration (MPI) model's structure in order to provide an intuitive explanation of why the model succeeds. We propose that the perception of the MBPI reflects a mechanism

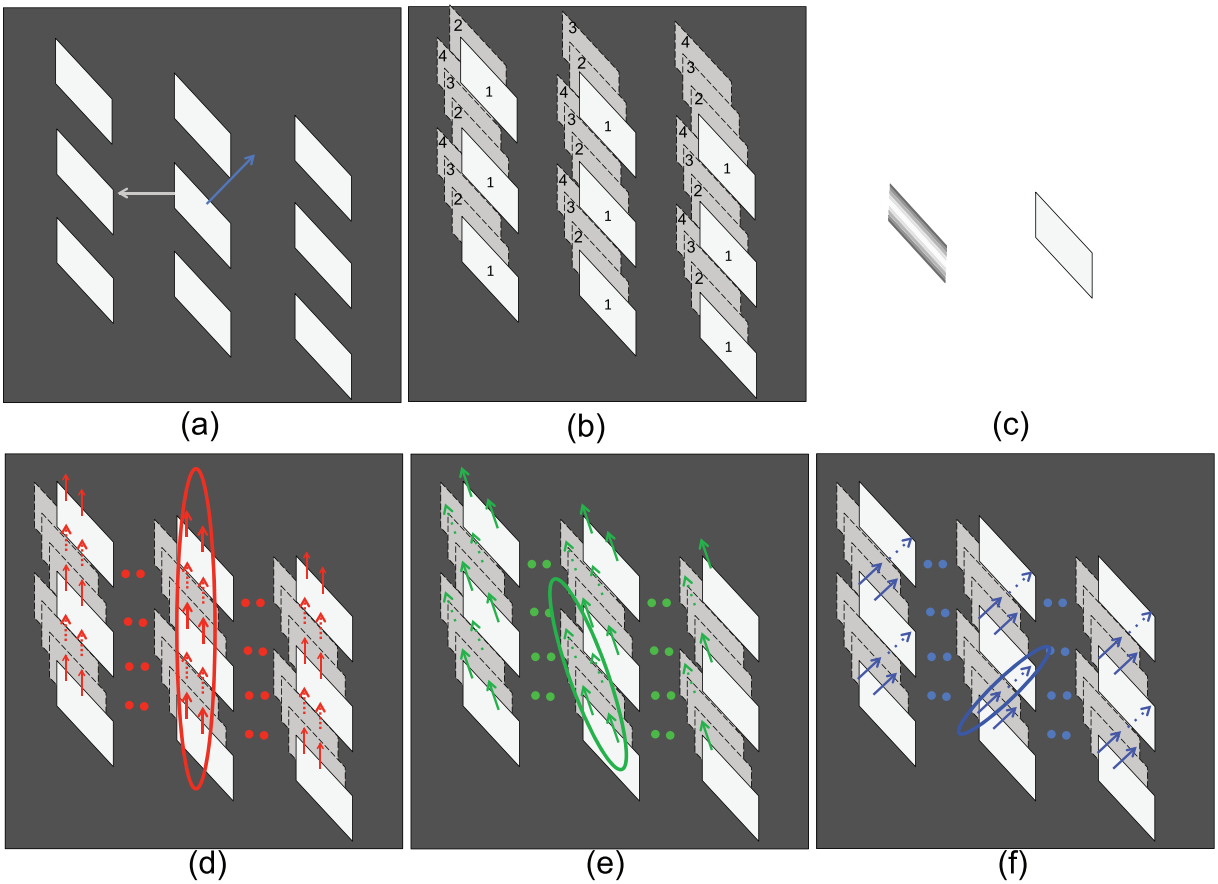

Figure 15. Examples of local motion energy traces for a moving-barber-pole stimulus in which the modulator moves to the left at 2.5 Hz and the carrier moves up to the right at 10 Hz. The carrier and the modulator are represented as square-waves with the dark regions set equal to the background. (a) Snapshot of a single frame. (b) Four consecutive frames shown together. For clarity, features in frames 2, 3, and 4 are shown in lower contrast. Bars with the same number belong to the same frame in the image sequence. (c) For clarity, the carrier is represented as square-waves. (d–f) Illustrations of local motion energies along the barber-pole direction, the feature direction, and the carrier motion direction, respectively. In each figure, spots represent points where local motion energies are zero. Solid arrows indicate motion energies between frames 1 and 2; dashed arrows indicate motion energies between frames 3 and 4.

that operates over a relatively large spatial area and prolonged time period. In Bowns' (2011) model, a moving feature leaves a trace. In the MPI model, the changing location of local motion energy leaves a trace. Like feature traces, local motion energy traces for a given image sequence are functions of both space and time. Consider for example a moving barber-pole stimulus, our standard stimulus, in which the modulator moves at a relatively slow temporal frequency oppositely to the carrier (Figure 15a). To simplify Figure 15, the sine-wave carrier and modulator are both represented as square-waves. The feature extraction process (described in detail below) produces the bright regions of the carrier square-wave (the bright regions in Figure 15). These features move in a consistent direction (i.e., the feature direction) as a result of the carrier's and modulator's cooperative movement. Figure 15b shows the locations of those features across four frames. Alternatively, Figure 15 can be thought of as the 2-D projection (i.e., the 2-D retinal image) of the three-dimensional (3-D) spatial-temporal feature function over four frames. In Figure 15d through f, the local motion energy traces are shown based on this 2-D projection of the 3-D feature functions.

In a 2-D motion signal, local motion energy at a point exists in many directions. In Figure 15 we show motion energies, represented as vectors, along three directions: the barber-pole direction (Figure 15d), the feature direction (Figure 15e) and the carrier motion direction (Figure 15f). In each of Figure 15d, e, and f, dots (instead of vectors) represent points with zero local motion energy in the corresponding direction. Regions outside of the barber-pole windows (i.e., zero contrast regions) contain mostly zero motion energies in all directions. The solid arrows indicate the motion

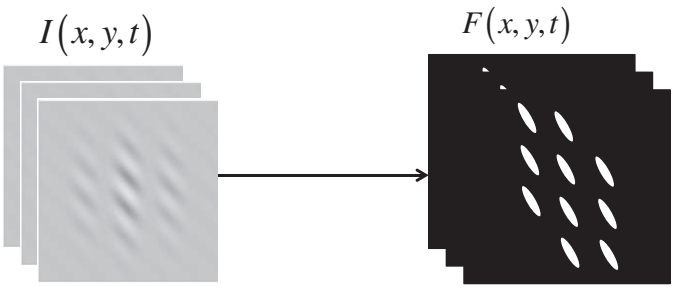

Figure 16. Feature extraction stage in the model.

energies computed between frames 1 and 2; the dashed arrows indicate the motion energies computed between frame 3 and 4. Motion energies between frame 2 and 3 are similar but are not drawn to reduce clutter. Motion energies for different frame pairs are labelled distinctively to reflect the fact that local motion energies vary with time. Although they are drawn in one figure panel, they actually appear in different frames in time. In this example, along the barber-pole direction (Figure 15e), both solid and dashed arrows connect to one another to form a consistent vertical pattern. Along the other two directions (Figure 15e, f), the vector connections are interrupted by zero motion energy regions. Therefore a mechanism that integrates local motion energies over an elongated region and over limited time period (i.e., one cycle of a 10 Hz stimulus) can potentially explain the peripheral MBPI in this stimulus. It is worth noting that the advantage of a vertical connection over other connections disappears when motion energies are traced for longer time period. The time period included in the integration is critical. If the time period included in the feature trace were to increase significantly, the vertical direction of Figure 15d would lose its advantage to the feature tracking direction of Figure 15e. The temporal coherence theory (Grzywacz et al., 1995) implicitly integrates over a large number of frames and therefore predicts that the feature (not the MBPI) direction would be perceived in this stimulus.

Model implementation

The MPI model uses some elements of the temporal coherence theory (Grzywacz et al., 1995) and the formotion model (Grossberg et al., 2001). The model assumes that the MBPI motion is mediated by a mechanism operating on the output of a local Fourier-energy motion system. A second stage process then connects similarly tuned local motion energy detectors along their preferred directions. We describe each step of the MPI model in more detail below. Figure 16 shows the result of applying this transducer to a moving barber pole stimulus.

Feature extraction transducer

The first-order motion system is approximately linear when the stimulus is of low contrast (Lu & Sperling, 1995; Van Santen & Sperling, 1984). In unpublished experiments, we found that the moving-barber-pole stimulus produced very different motion perceptions at low versus high contrasts. The low-contrast data are predictable with a linear transducer. The high contrast data described here require a highly nonlinear transducer, essentially a half-wave rectifier. Data predictions were possible with either a negative half-wave rectifier that emphasizes the dark phase of sine wave, or a positive half-wave rectifier that emphasizes the light phase of the sine waves or, obviously, with both systems working independently.

Various forms of rectification were tried, and the model's predictions were very insensitive to the form. Here we chose, arbitrarily, to model the transducer as a positive half-wave rectifier described in Equation 2 below. For high contrast stimuli, this stage creates a feature map in which the features are the light areas of the stimulus. Such a nonlinear, feature-extracting stage was necessary to account for the perceived feature-motion direction at high temporal frequencies.

The nonlinear transducer function has two positive parameters, an amplification factor a, and a threshold p. Let $I(x, y, t)$ be the image sequence. Then the output of the transducer, the extracted feature map, is given by Equation 2:

$$Ft(x, y, t) = \begin{cases} aI(x, y, t) & \text{if} I(x, y, t) \geq p \\ I(x, y, t) & \text{if} I(x, y, t) < p \end{cases} \quad (2)$$

where p was 20% and a was 5.0. The exact values of a and p are not critical as long as features (e.g., the light bars) are well differentiated from the background.

First-order local motion detector

An Elaborated Reichardt Detector (ERD; Van Santen & Sperling, 1984, 1985) with a critical modification of the shape of the spatial filters' receptive fields (Figure 17a) detects local motion energies after the transducer distortion. A typical ERD has two input spatial filters occupying the same area in the visual field. Typically, the two spatial filters have receptive fields elongated in the direction that is perpendicular to the detector's preferred direction of motion. In this paper we call this type of configuration an orthogonal design (ORTH). Physiological support for the ORTH comes from studies of directionally selective cells in V1 that use the reverse correlation technique (DeAngelis, Ohzawa, & Freeman, 1993, 1995; De Valois, Yund, & Hepler, 1982; Hubel & Wiesel, 1962). Psychophysical support for ORTH comes from a study of the contrast threshold for motion detection (Anderson & Burr,

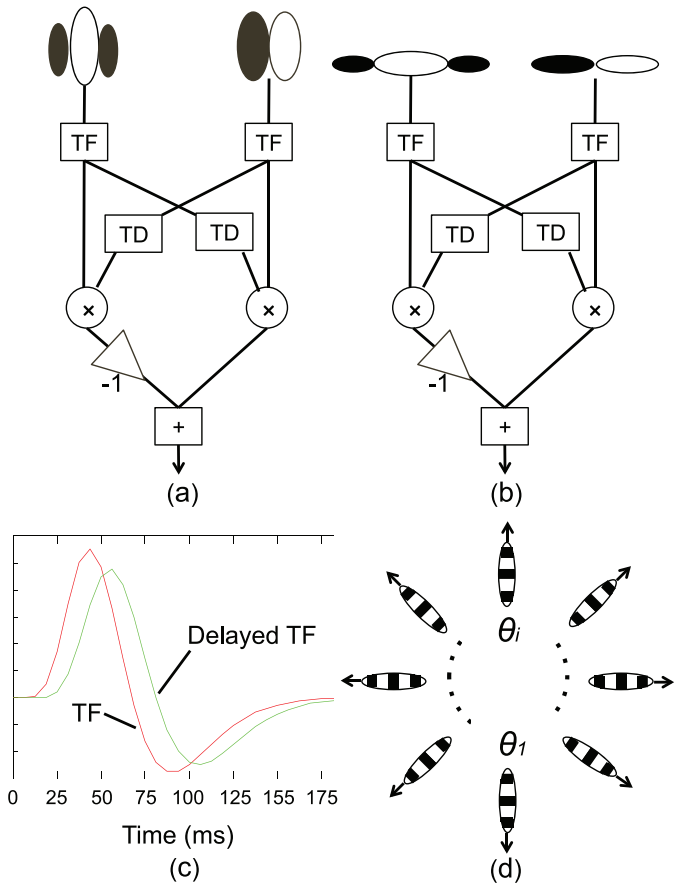

Figure 17. Diagrams of the local motion detector components of the MPI model. (a) A typical ERD. Spatial filters are elongated perpendicularly to the detector's preferred motion direction (Van Santen & Sperling, 1984, 1985). (b) The ERD used in the MPI model has spatial filters extended in the detector's preferred direction (EXT). In (a) and (b), detector's preferred motion direction is horizontal, with positive outputs indicating rightward motion. (c) The impulse response of the temporal filter TF and the impulse response after passing through TD, *TF* × *TD*. (d) Illustration of the range of directions to which ERDs are tuned. The striped ovals indicate ERDs as illustrated in (b) tuned to the directions of the arrows. The full range of motion directions is assumed to be represented in every local area in the region of interest.

1991). However, evidence also exists for a different receptive field shape in which the spatial receptive field is elongated in the same direction as the motion detector's preferred direction (Fredericksen et al., 1994a, 1994b; Geisler et al., 2001; Jancke, 2000; Van Doorn & Koenderink, 1984). We call this type of configuration an extended design (EXT). A recent study shows that EXT can help in detecting prolonged motion such as motion streaks (as opposed to short-lived, brief motion; Pavan et al., 2011). The evidence suggests that two quite different receptive field configurations both exist but may be differentially effective depending on stimulus contrast and other factors.[1] To account for our data, it was necessary to use a EXT design for our ERD's spatial receptive field; based on Fredericksen et al. (1994b) and Van Doorn and Koenderink (1984), we chose an aspect ratio of 10 : 1. Finally, the bandwidth of the spatial filter was fixed at 1 octave. Defining the bandwidth in terms of octaves means that the configuration of the spatial filter scales with the filter's optimal spatial frequency.

In addition to the spatial filters, the ERD has two temporal filters. The first temporal filter (TF in Figure 17b) reflects the low- and band-pass characteristic of human observers' temporal sensitivity (Robson, 1966; Watson & Ahumada, 1985). The impulse function is depicted by the red curve in Figure 17b. The second temporal filter (TD in Figure 17b) is a first-order, low-pass filter with an impulse response defined by $e^{-t/\tau}$ for $t \geq 0$, and 0 for $t < 0$. TD delays its input and this delayed input is compared with the undelayed input in the multiplier unit, ×. The impulse function of the temporal filter combining the first and second temporal filters is depicted by the green curve in Figure 17b. This configuration and the characteristics of the temporal filters are largely consistent with conventional models of motion energy detectors (Adelson & Bergen, 1985; Van Santen & Sperling, 1985; Watson & Ahumada, 1985).

The responses of the ERDs are normalized by the size of their receptive fields (not shown in Figure 17) to ensure that all ERDs, independent of their spatial frequency tuning, respond with equal intensity when presented with their optimal stimulus. For a given direction, the MPI model considers only the ERD optimally tuned to the spatial frequency along that given direction. To find the optimal spatial frequency in the moving-barber-pole stimuli for every direction, we first calculated the optimal spatial frequencies analytically. Then a computer simulation was conducted to ensure that ERDs with those optimal spatial frequencies indeed produced the strongest response. To characterize the first-stage operation, we define $ERD_{\theta i}$ as the ERD that is tuned in the direction θ_i and that contains optimally tuned spatial filters in the direction θ_i. The output of the Local Motion Energy Detector stage is a field of local motion energies $E_i(x, y, t)$ in the direction θ_i:

$$E_i(x, y, t) = ERD_{\theta_i}(Ft(x, y, t)) \quad (3)$$

The MPI model assumes a second stage mechanism that integrates the output from the Local Motion Energy Detector over a relatively large area at a limited time interval. Each component unit of the second-stage mechanism has its own directional tuning property (Figure 18a). It aggregates half-rectified responses from an array of local ERDs that are tuned to the same

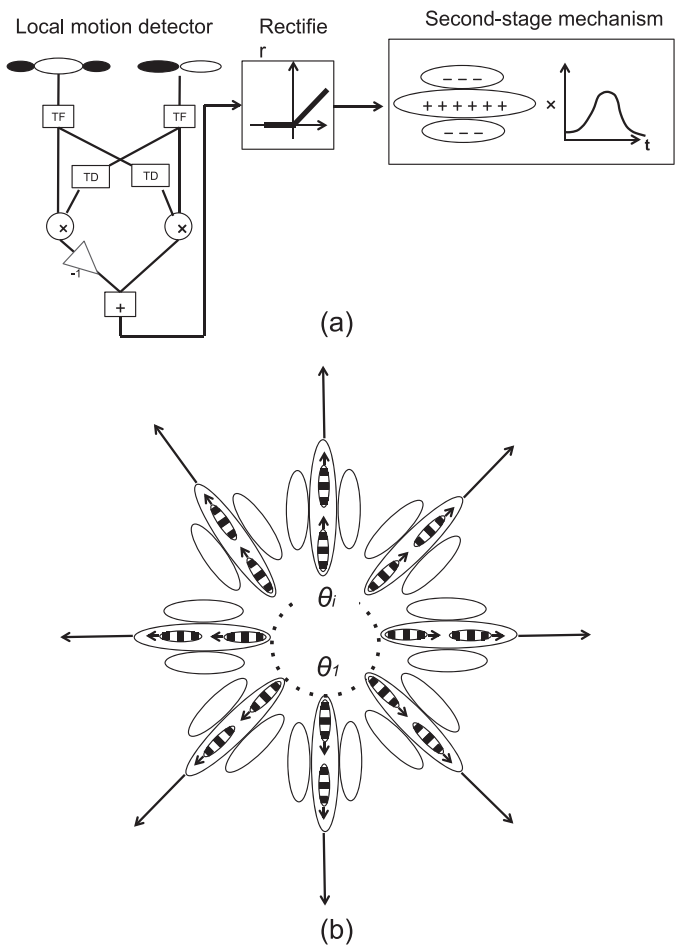

Figure 18. Diagrams for the second-stage mechanism. (a) A directionally tuned, higher-order mechanism aggregates half-wave-rectified responses of local motion energy detectors. The aggregation is carried out along a path that is inline with the higher-order mechanism's preferred direction. Two weak inhibitory regions are added flanking the aggregation region to improve the mechanism's directional selectivity. (b) Higher-order mechanisms tuned to various directions.

direction as the second stage mechanism. This aggregation is carried out over a limited time period σ along a spatial path that is in the common preferred direction θ_i of the MPI model's first- and second-stage mechanisms. To improve the directional selectivity of the second-stage mechanism, a lateral inhibition region that flanks the spatial integration path is included. Together, the integration and inhibition area form a higher-order receptive field for the second stage mechanism. We implemented the integration process by applying a 3-D Gabor filter to the local motion energies (Equation 3) computed by the previous stage. To describe the second-state integration we need to define the following:

First, let $G_i(x, y, t)$ be a 3-D Gabor filter that aggregates local motion energies at a point x, y and time t a long the direction θ_i

$$G_i(x,y,t) = \exp\left(-\frac{(x\cos\theta_i + y\sin\theta_i)^2}{2\sigma_x^2} - \frac{(y\cos\theta_i - x\sin\theta_i)^2}{2\left(\frac{\sigma_x}{\alpha}\right)^2}\right)\cos(2\pi\phi(x\cos\theta_i + y\sin\theta_i))\exp\left(-\frac{t^2}{2\sigma_t^2}\right) \tag{4}$$

The 3-D Gabor filter $G_i(x, y, t)$ can be thought of as a stack of spatial 2-D Gabor filters whose amplitudes are modulated by a temporal Gaussian function $\exp(-\frac{t^2}{2\sigma_t^2})$.

Then let $L_i(x, y, t)$ be a weighed sum of the local motion energies $E_i(x, y, t)$ along the direction θ_i

$$L_i(x,y,t) = \iiint G_i(x-\gamma, y-\eta, t-\tau)E_i(\gamma,\eta,\tau)\mathrm{d}\gamma\mathrm{d}\eta\mathrm{d}\tau \tag{5}$$

The response, RES_i, of a second-stage mechanism with a preferred direction of θ_i, is given by the variance, i.e., the power, of $L_i(x, y, t)$ that can be understood as a measure of response amplitude of the higher-order receptive field. RES_i gives the total power in the direction θ_i over the whole stimulus for its entire duration.

$$RES_i = Var(L_i(x,y,t)) \tag{6}$$

Finally the predicted motion direction PMD is given by the mean of the responses RES_i over all simulated directions.

$$PMD = \frac{\sum_{i=1}^{n}\theta_i RES_i}{\sum_{i=1}^{n} RES_i} \tag{7}$$

The full MPI model is outlined in Figure 19.

Explanation of model parameters: The model has three free parameters, all concerning the characteristic of the second-stage mechanism. The first free parameter α is the aspect ratio of the 3-D Gabor filter that is used to implement the spatial-temporal motion energy integration. α controls the elongation of the higher-order receptive field. A more elongated higher-order receptive field extends the spatial integration area, hence promoting detection of barber-pole motion. The second free parameter σ_t controls the temporal integration window. Increasing σ_t increases the model's feature tracking capability.

The parameter φ controls lateral inhibition. Lateral inhibition in motion integration is similar to the optimal

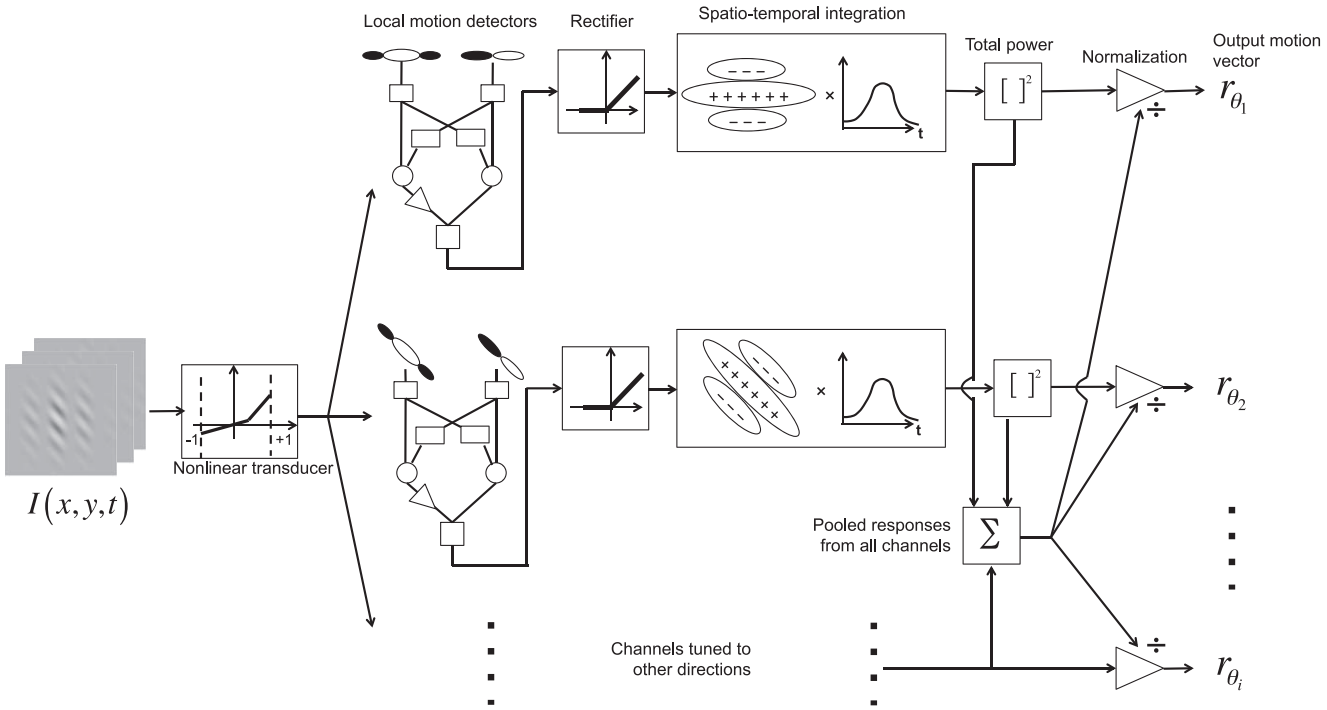

Figure 19. Functional flowchart for the MPI model. In the initial nonlinear process, all input image intensities that exceed a certain threshold are amplified. The resultant image sequence is processed by an ensemble of special motion channels tuned to a full range of directions. Within each directionally tuned special channel, local motion energies are computed by a spatially distributed array of ERDs whose spatial input units are elongated in the direction of the ERDs' preferred motion direction. The output of each ERD is half-wave rectified; thereby the output is positive only in the preferred and not the anti-preferred direction. The rectified output of an array of similar ERDs is integrated along a spatial path that is parallel with the ERDs' preferred direction, i.e., the outputs are passed through a low-pass temporal filter and a 2-D spatial filter whose profile matches the spatial integration path. The motion-power of each directionally tuned channel is normalized by the total power of all channels. A subsequent process (not shown) computes the mean direction of the normalized vector of directional motion outputs.

integration area for shearing motion (Golomb, Andersen, Nakayama, MacLeod, & Wong, 1985). In a typical shearing motion, the velocity of the moving elements varies as a function of positions along the axis perpendicular to the direction of motion. In one experiment in Golomb et al.'s (1985) study, the shearing motion was made of a field of random dots moving horizontally. The velocity within each row was constant but varied vertically following a sinusoidal function of vertical positions. Dots in different rows could move in opposite directions and some rows could be completely static. It was reported that the detectability of this kind of shearing motion varied with the spatial frequency of the vertical sinusoidal function, peaking at about 0.6–0.7 c/d. Modulations of lower spatial frequency impaired the detection of the vertical sinusoid modulation, indicating the existence of an optimal integration area being flanked by a lateral inhibition region.

Model prediction

To derive predictions of our data, we first consider only the moving-barber-pole stimuli of 10 Hz carrier temporal frequency with the seven different modulator frequencies. The two stage processes constitute a compound mechanism that is directionally tuned (Figure 18b). We applied the two-stage mechanisms tuned to 36 different directions. Because of computational limitations, a coarse grid search (seven days) was used to find approximately optimal values for φ, α, and σ_t. Interestingly, all parameters have great degrees of tolerance; when φ varies between 0.8 c/d and 1.2 c/d, σ_t varies between 40 ms and 80 ms and α varies between 0.3 and 1.2, the accounted data variance only varies between 80% and 92%. The semi-optimal parameter range is summarized in Table 1.

Figure 20a shows model responses RES_i as functions of the preferred direction θ_i, produced by a model implementation with α set to 0.5, σ_t set to 60ms and φ set to 0.8 c/d. We deliberately choose to use the lowest possible value for φ to be consistent with the optimal integration area found in Golomb et al. (1985). Figure 20b shows the predicted motion direction (solid black line) for each of the seven modulator temporal frequencies. With the choice of these three values, the model accounts for 90% of the variance in the data. The

Free parameters	Values
ϕ	0.8∼1.2 c/d
α	0.3∼1.2
σ_t	40∼60 ms

Table 1. Free parameters and their optimal values for the moving-barber-pole stimulus with carrier temporal frequency of 10 Hz. *Notes*: The parameter ϕ controls lateral inhibition. α controls the elongation of the higher-order receptive field. σ_t controls the temporal integration window.

model describes the essential features of these data: perception of the rigid motion direction when the modulator moves rapidly either with or opposed to the carrier motion, and the barber-pole motion direction in between.

To find the optimal parameter set for other carrier temporal frequencies, in principle one should run simulations with all possible parameters to each human data set, which is very computationally intensive. However we find that the parameter set obtained for the 10 Hz carrier temporal frequency condition automatically fits other carrier temporal frequencies well. Figure 21 gives examples of the MPI model predictions for stimuli with carrier temporal frequency of 10, 5, 2.5, and 0 Hz. The model prediction successfully captures the transition of the MBPI motion to the rigid motion as the carrier temporal frequency decreases. Using the same parameter values as mentioned above, the total accounted variance for stimuli of four different carrier frequencies is 97%.

Discussion: Isolating a single mechanism?

The novel moving-barber-pole stimulus allows us to discriminate between various motion computations because of the different predictions they make for the perceived direction of the stimulus motion. Both peripheral viewing of the stimulus and high temporal frequencies reduce the contributions of higher-order motion computations (Lu & Sperling, 1995, 1999) and of other possible complex mechanisms. Ideally, this impoverished stimulus isolates a single mechanism that is responsible for the observed MBPI. The MPI model was developed to explain the computations of this mechanism. The MPI model successfully accounted for the data from the full range of peripherally viewed stimuli, and also for foveally viewed stimuli when higher-order motion was cancelled. In the special case of a stationary modulator, in the peripheral viewing conditions explored in this study, the MPI model also accounted for the classic BPI. In central vision (Figure 11) and at low or intermediate temporal frequencies, other mechanisms came into play.

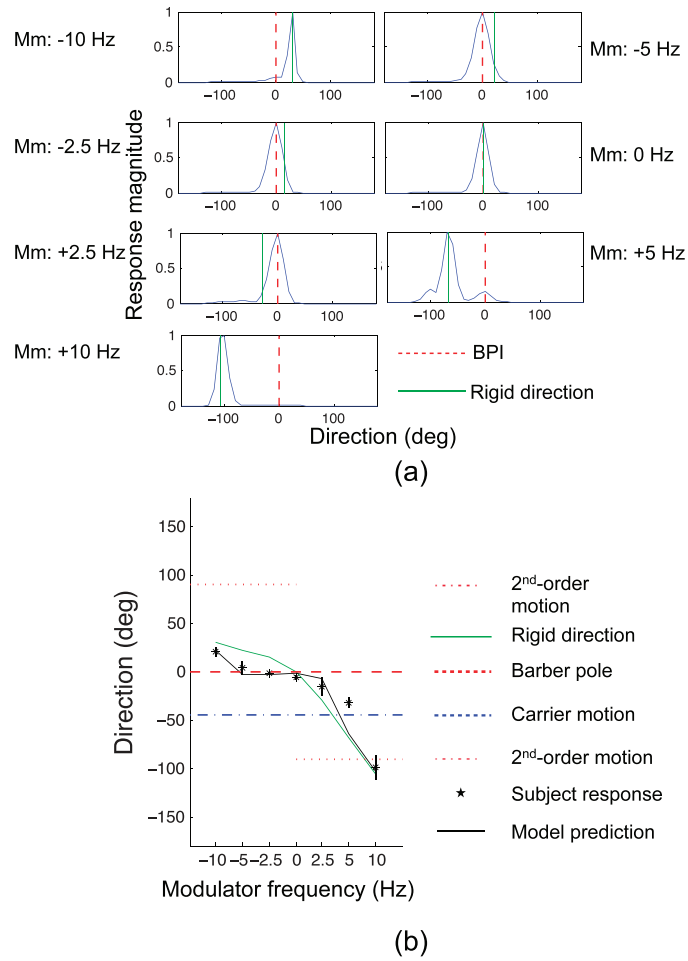

Figure 20. Model predictions for moving-barber-pole stimuli of seven modulator temporal frequencies. The carrier temporal frequency is always 10 Hz. (a) A compound, directionally tuned mechanism's response as functions of the mechanisms' tuned direction. Each panel corresponds to a stimulus of a particular modulator temporal frequency presented in the same format as in Figure 5a. The green solid line indicates the rigid direction and the dashed red line indicates the barber-pole direction. (b) The predicted motion direction as a function of the modulator temporal frequency. Each prediction value is obtained by finding the mean of the response distribution curve in each panel in (a). Asterisks represent the mean of the four subjects' responses. Error bars represent 95% confidence intervals. Predicted motion directions account for 90% of the data variance.

The MPI model proposes a regime in which locally distributed, low-level motion sensitive neurons are connected to a higher-level processing stage that aggregates motion energies in a particular direction within a defined time period. Somewhat similar neural connections to extract object contours from common motion signals were proposed by Ledgeway and Hess

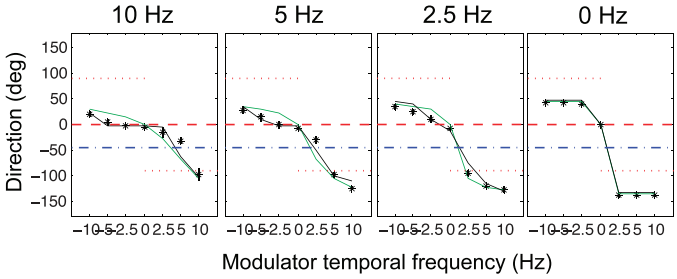

Figure 21. Model predictions for the moving-barber-pole stimuli in which the temporal frequency of the carrier varies from 10 Hz (the left-most column) to 0 Hz (the right-most column). The model prediction accounts for 97% of the total data variance.

(2002, 2006) and by Watamaniuk, McKee, and Grzywacz (1995) to extract extended motion paths.

Many higher-order motion computations such as the computation of heading direction and of 3-D structure from 2-D motion are based on initial low-level motion computations that critically involve velocity. The MPI model specifically uses only directional motion energy and does not explicitly involve velocity.

Summary and conclusions

The moving-barber-pole display

The normal BPI is a phenomenon that reflects the influence of form (the window through which a grating is perceived) on the perception of motion (the direction in which the grating appears to move). To better understand the mechanism underlying the form-motion interaction, the current study introduced a novel moving-barber-pole display. The moving-barber-pole display has a translating, rather than a static, modulator window. Therefore, the rigid direction (i.e., the directions of 2-D spatial features such as end-stops), is different from the orientation of the modulator window. The moving-barber-pole display effectively differentiates three different motion directions: the rigid motion direction (which includes the direction of end stops and of other features), the motion direction of the main Fourier components, and motion in the direction of an elongated modulator window.

The MBPI

Displays were viewed peripherally and at high temporal frequencies in order to minimize the influence of higher-order motion computations and thereby to isolate a single motion mechanism. In peripheral viewing, three of four subjects reported strong MBPI (i.e., the perceived motion direction was in line with the orientation of the modulator window), the fourth subject's perceived direction was slightly deviated. When masking signals were added in between the modulator windows, the MBPI was retained by the other three. In the fovea, perceived motion was more complicated, due at least in part to concurrent perception of the higher-order modulator motion that was not visible in the periphery. Adding a masking pattern to interfere with the higher-order modulator motion produced a pure MBPI in the fovea. That reducing the subjects' ability to perceive the modulator's orientation strengthens the MBPI means that the MBPI does not utilize the perceived modulator orientation, which in turn means that motion-streak models do not apply to the MBPI. Based on these observations, we conclude that the BPI generated by the moving-barber-pole stimuli was neither a result of feature tracking nor a combination of the first-order carrier motion with the higher-order modulator motion, nor a combination of perceived orientation of the modulator with the carrier motion.

Two other results: In peripheral viewing, reducing the carrier temporal frequency shifted the perceived motion direction away from the barber-pole direction and ultimately in the direction of the rigid motion. How the transition between the rigid motion direction (feature motion direction) and the MBPI depends on carrier and modulator temporal frequencies is an important feature of the experimental results that potentially can be used to discriminate between theories. Changing the phase of motion in adjacent barber poles had no effect on the MBPI, showing that, at an early stage, the MBPI is computed within a barber pole, not between barber poles.

The MPI model

A three-stage MPI model was proposed. First, local motion energies are computed by ERDs whose spatial input units are elongated in the direction of the ERD's preferred motion direction. Unlike typical motion models, the ERD's motion receptive fields are elongated parallel with rather than perpendicular to the motion direction (EXT design). The EXT design is critical for the predictions. Then, the local motion energies are integrated along linear spatial paths (in all directions) and over a time period of 40–60 ms. The precise length of the path was not critical in the model. Finally, the predicted motion direction is given by the vector mean of the responses over all directions. The MPI model embodies entirely feed-forward computations. It accounts for the large variety of observed BPI phenomena entirely within the motion system and does

not require any information computed in shape or form systems.

It was not technically feasible to estimate optimal parameters for the MPI model. Even with merely three "good" parameters (temporal integration period, path length, and path spatial frequency), the model accounts with great accuracy for a variety of very different data, i.e., the change from MBPI to rigid (feature) motion direction with carrier temporal frequency, the interaction of modulator and carrier temporal frequencies, and the dependence on local versus global stimulus factors.

Keywords: 2D motion, form-motion-interaction, Barber-Pole-Illusion, motion integration, motion model

Acknowledgements

This work was supported by NSF Award BCS-0843897. We thank the two anonymous reviewers for their suggestions that helped to greatly improve the quality of the manuscript.

Commercial relationships: none.
Corresponding author: Peng Sun.
Email: peng.sun@uci.edu.
Address: Department of Cognitive Sciences, University of California Irvine, Irvine, CA, USA.

Footnote

[1]Although this topic is beyond the scope of the current study, we would like to point out that the receptive field shape is found to vary as stimulus contrast changes (Lombrozo et al., 2005). Thus the receptive field shape derived from the contrast threshold of motion detection (e.g., Anderson & Burr, 1991) is not necessarily the shape of high-contrast receptive fields.

References

Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, *2*(2), 284–299.

Adelson, E. H., & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, *300*(5892), 523–525.

Anderson, S. J., & Burr, D. C. (1991). Spatial summation properties of directionally selective mechanisms in human vision. *Journal of the Optical Society of America A*, *8*(8), 1330–1339.

Badcock, D. R., & Dickinson, J. E. (2009). Second-order orientation cues to the axis of motion. *Vision Research*, *49*(3), 407–415.

Badcock, D. R., McKendrick, A. M., & Ma-Wyatt, A. (2003). Pattern cues disambiguate perceived direction in simple moving stimuli. *Vision Research*, *43*(22), 2291–2301.

Barlow, H. B., & Olshausen, B. A. (2004). Convergent evidence for the visual analysis of optic flow through anisotropic attenuation of high spatial frequencies. *Journal of Vision*, *4*(6):1, 415–426, http://www.journalofvision.org/content/4/6/1, doi: 10.1167/4.6.1. [PubMed] [Article]

Berens, P. (2009). Circstat: A Matlab toolbox for circular statistics. *Journal of Statistical Software*, *31*(10), 1–21.

Beutter, B. R., Mulligan, J. B., & Stone, L. S. (1996). The barberplaid illusion: Plaid motion is biased by elongated apertures. *Vision Research*, *36*(19), 3061–3075.

Bowns, L. (1996). Evidence for a feature tracking explanation of why type II plaids move in the vector sum direction at short durations. *Vision Research*, *36*(22), 3685–3694.

Bowns, L. (2011). Taking the energy out of spatio-temporal energy models of human motion processing: The component level feature model. *Vision Research*, *51*(23-24), 2425–2430.

Brainard, D. H. (1997). The Pscyhophysics Toolbox. *Spatial Vision*, *10*(4), 433–436.

Burr, D. C., & Ross, J. (2002). Direct evidence that speedlines influence motion mechanisms. *Journal of Neuroscience*, *22*(19), 8661–8664.

Castet, E., Charton, V., & Dufour, A. (1999). The extrinsic/intrinsic classification of two-dimensional motion signals with barber-pole stimuli. *Vision Research*, *39*(5), 915–932.

Castet, E., & Wuerger, S. (1997). Perception of moving lines: Interactions between local perpendicular signals and 2D motion signals. *Vision Research*, *37*(6), 705–720.

Chubb, C., & Sperling, G. (1989). Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proceedings of the National Academy of Sciences*, *86*(8), 2985–2989.

DeAngelis, G. C., Ohzawa, I., & Freeman, R. (1993). Spatiotemporal organization of simple-cell receptive fields in the cat's striate cortex. I. General characteristics and postnatal development. *Journal of Neurophysiology*, *69*(4), 1091–1117.

DeAngelis, G. C., Ohzawa, I., & Freeman, R. D.

(1995). Receptive-field dynamics in the central visual pathways. *Trends in Neurosciences*, *18*(10), 451–458.

De Valois, R. L., Yund, E. W., & Hepler, N. (1982). The orientation and direction selectivity of cells in macaque visual cortex. *Vision Research*, *22*(5), 531–544.

DeYoe, E. A., & Van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in Neurosciences*, *11*(5), 219–226.

Edwards, M., & Crane, M. F. (2007). Motion streaks improve motion detection. *Vision Research*, *47*(6), 828–833.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local association field. *Vision Research*, *33*(2), 173–193.

Fisher, N., & Zanker, J. M. (2001). The directional tuning of the barber-pole illusion. *Perception-London*, *30*(11), 1321–1336.

Fredericksen, R., Verstraten, F., & Van De Grind, W. (1994a). An analysis of the temporal integration mechanism in human motion perception. *Vision Research*, *34*(23), 3153–3170.

Fredericksen, R., Verstraten, F., & Van De Grind, W. (1994b). Spatial summation and its interaction with the temporal integration mechanism in human motion perception. *Vision Research*, *34*(23), 3171–3188.

Geisler, W. S. (1999). Motion streaks provide a spatial code for motion direction. *Nature*, *400*(6739), 65–68.

Geisler, W. S., Albrecht, D. G., Crane, A. M., & Stern, L. (2001). Motion direction signals in the primary visual cortex of cat and monkey. *Visual Neuroscience*, *18*(4), 501–516.

Golomb, B., Andersen, R., Nakayama, K., MacLeod, D., & Wong, A. (1985). Visual thresholds for shearing motion in monkey and man. *Vision Research*, *25*(6), 813–820.

Grossberg, S., Mingolla, E., & Viswanathan, L. (2001). Neural dynamics of motion integration and segmentation within and across apertures. *Vision Research*, *41*(19), 2521–2553.

Grzywacz, N. M., Watamaniuk, S. N., & Mckee, S. P. (1995). Temporal coherence theory for the detection and measurement of visual motion. *Vision Research*, *35*(22), 3183–3203.

Heeger, D. J. (1987). Model for the extraction of image flow. *Journal of the Optical Society of America A*, *4*(8), 1455–1471.

Hubel, D. H., & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology*, *160*(1), 106–154.

Jancke, D. (2000). Orientation formed by a spot's trajectory: A two-dimensional population approach in primary visual cortex. *Journal of Neuroscience*, *20*(14), 1–6.

Kooi, F. L. (1993). Local direction of edge motion causes and abolishes the barberpole illusion. *Vision Research*, *33*(16), 2347–2351.

Kourtzi, Z., & Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *Journal of Cognitive Neuroscience*, *12*(1), 48–55.

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

Lalanne, C., & Lorenceau, J. (2006). Directional shifts in the barber pole illusion: Effects of spatial frequency, spatial adaptation, and lateral masking. *Visual Neuroscience*, *23*(5), 729–739.

Ledgeway, T., & Hess, R. F. (2002). Rules for combining the outputs of local motion detectors to define simple contours. *Vision Research*, *42*(5), 653–659.

Ledgeway, T., & Hess, R. F. (2006). The spatial frequency and orientation selectivity of the mechanisms that extract motion-defined contours. *Vision Research*, *46*(4), 568–578.

Lidén, L., & Mingolla, E. (1998). Monocular occlusion cues alter the influence of terminator motion in the barber pole phenomenon. *Vision Research*, *38*(24), 3883–3898.

Lombrozo, T., Judson, J., MacLeod, D. I., Lombrozo, T., Judson, J., & MacLeod, D. I. (2005). Flexibility of spatial averaging in visual perception. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1564), 725–732.

Lorenceau, J., Shiffrar, M., Wells, N., & Castet, E. (1993). Different motion sensitive units are involved in recovering the direction of moving lines. *Vision Research*, *33*(9), 1207–1217.

Lu, Z.-L., & Sperling, G. (1995). The functional architecture of human visual motion perception. *Vision Research*, *35*(19), 2697–2722.

Lu, Z.-L., & Sperling, G. (1999). Second-order reversed phi. *Perception & Psychophysics*, *61*(6), 1075–1088.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: Henry Holt and Co.

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

Mather, G., Pavan, A., Bellacosa, R. M., & Casco, C.

(2012). Psychophysical evidence for interactions between visual motion and form processing at the level of motion integrating receptive fields. *Neuropsychologia*, *50*(1), 153–159.

Mather, G., Pavan, A., Marotti, R. B., Campana, G., & Casco, C. (2013). Interactions between motion and form processing in the human visual system. *Frontiers in Computational Neuroscience*, *7*(65), 1–6.

Maunsell, J. H., & Newsome, W. T. (1987). Visual processing in monkey extrastriate cortex. *Annual Review of Neuroscience*, *10*(1), 363–401.

Pack, C. C., Gartland, A. J., & Born, R. T. (2004). Integration of contour and terminator signals in visual area MT of alert macaque. *Journal of Neuroscience*, *24*(13), 3268–3280.

Pack, C. C., Livingstone, M. S., Duffy, K. R., & Born, R. T. (2003). End-stopping and the aperture problem: Two-dimensional motion signals in macaque V1. *Neuron*, *39*(4), 671–680.

Pavan, A., Casco, C., Mather, G., Bellacosa, R. M., Cuturi, L. F., & Campana, G. (2011). The effect of spatial orientation on detecting motion trajectories in noise. *Vision Research*, *51*(18), 2077–2084.

Pavan, A., Marotti, R. B., & Mather, G. (2013). Motion-form interactions beyond the motion integration level: Evidence for interactions between orientation and optic flow signals. *Journal of Vision*, *13*(6):16, 1–13, http://www.journalofvision.org/content/13/6/16, doi:10.1167/13.6.16. [PubMed] [Article]

Perrone, J. A. (2012). A neural-based code for computing image velocity from small sets of middle temporal (MT/V5) neuron inputs. *Journal of Vision*, *12*(8):1, 1–31, http://www.journalofvision.org/content/12/8/1, doi:10.1167/12.8.1. [PubMed] [Article]

Perrone, J. A. (2004). A visual motion sensor based on the properties of V1 and MT neurons. *Vision Research*, *44*(15), 1733–1755.

Robson, J. (1966). Spatial and temporal contrast-sensitivity functions of the visual system. *Journal of the Optical Society of America A*, *56*(8), 1141–1142.

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

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

Schrater, P. R., Knill, D. C., & Simoncelli, E. P. (2000). Mechanisms of visual motion detection. *Nature Neuroscience*, *3*(1), 64–68.

Shimojo, S., Silverman, G. H., & Nakayama, K. (1989). Occlusion and the solution to the aperture problem for motion. *Vision Research*, *29*(5), 619–626.

Simoncelli, E. P., & Heeger, D. J. (1998). A model of neuronal responses in visual area MT. *Vision Research*, *38*(5), 743–761.

Sperling, G., & Liu, D. (2009). The vector sum of motion strengths describes the perceived motion direction of first-order plaids [abstract]. *Perception*, *38*, 57.

Tse, P., & Hsieh, P.-J. (2006). The infinite regress illusion reveals faulty integration of local and global motion signals. *Vision Research*, *46*(22), 3881–3885.

Tsui, J. M., Hunter, J. N., Born, R. T., & Pack, C. C. (2010). The role of V1 surround suppression in MT motion integration. *Journal of Neurophysiology*, *103*(6), 3123–3138.

Ungerleider, L., & Mishkin, M. (1982). Two cortical visual systems. In D. Ingle, M. Goodale, & R. Mansfield (Eds.), *Analysis of visual behavior analysis of visual behavior*. Cambridge, MA: MIT Press.

Van Doorn, A. J., & Koenderink, J. J. (1984). Spatiotemporal integration in the detection of coherent motion. *Vision Research*, *24*(1), 47–53.

Van Santen, J. P., & Sperling, G. (1985). Elaborated reichardt detectors. *Journal of the Optical Society of America A*, *2*(2), 300–320.

Van Santen, J. P., & Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America A*, *1*(5), 451–473.

Wallach, H. (1935). Über visuell wahrgenommene bewegungsrichtung [Translation: On the visually perceived direction of motion]. *Psychologische Forschung*, *20*(1), 325–380.

Watamaniuk, S. N., McKee, S. P., & Grzywacz, N. M. (1995). Detecting a trajectory embedded in random-direction motion noise. *Vision Research*, *35*(1), 65–77.

Watson, A. B., & Ahumada, A. J. (1985). Model of human visual-motion sensing. *Optical Society of America, Journal, A: Optics & Image Science*, *2,* 322–342.

Werkhoven, P., Sperling, G., & Chubb, C. (1993). The dimensionality of texture-defined motion: A single channel theory. *Vision Research*, *33*(4), 463–485.

Wilson, H., Ferrera, V., & Yo, C. (1992). A psychophysically motivated model for two dimensional motion perception. *Visual Neuroscience*, *9*(1), 79–97.