American Psychological Association

Psychological Review

ISSN: 0033-295X

http://dx.doi.org/10.1037/rev0000180

Theory of the Perceived Motion Direction of Equal-Spatial-Frequency Plaid Stimuli

AQ: 1

AQ: au
AQ: 2

George Sperling, Peng Sun, Dantian Liu, and Ling Lin
University of California, Irvine

AQ: 3 At an early stage, 3 different systems independently extract visual motion information from visual inputs. At later stages, these systems combine their outputs. Here, we consider a much studied (>650 publications) class of visual stimuli, plaids, which are combinations of 2 sine waves. Currently, there is no quantitative theory that can account for the perceived motion of plaids. We consider only perceived plaid direction, not speed, and obtain a large set of data exploring the various dimensions in which same-spatial-frequency plaids differ. We find that only 2 of the 3 motion systems are active in plaid AQ: 4 processing, and that plaids with temporal frequencies 10 Hz or greater typically stimulate only the first-order motion system, which combines the plaid components by vector summation: Each plaid component is represented by a contrast-strength vector whose length is contrast-squared times a factor representing the relative effectiveness of that component's temporal frequency. The third-order system, which becomes primary at low temporal frequencies, also represents a plaid as 2 vectors that sum according to their contrast strength: a pure plaid in which both components have equal contrast and a residual sine wave. Second-order motion is irrelevant for these plaids. These principles enable a AQ: 5 contrast-strength-vector summation theory for the responses of the first-order and third-order motion systems. With zero parameters estimated from the data, the theory captures the essence of the full range of the plaid data and supports the counterintuitive hypothesis that motion direction is processed independently of speed at early stages of visual processing.

Keywords: motion perception, plaids, theory, sine waves, motion systems

AQ: 6 *Supplemental materials:* http://dx.doi.org/10.1037/rev0000180.supp

Because sine waves are basis functions in linear systems analysis, a great deal of research on visual motion perception has been AQ: 7 focused on the perception of moving sine-wave gratings (e.g., Burr & Thompson, 2011; Kelly, 1979). However, even though real-world stimuli can be considered as being composed of sine waves, pure sine-wave visual gratings seldom occur outside of the laboratory. Therefore, an obvious next step in the systematic analysis of visual motion perception was to consider the motion of combinations of sine waves, beginning with stimuli composed of just two sine waves. A stimulus composed of two superimposed sine-wave AQ: 8 gratings moving in independent directions has been called a *plaid* (Adelson & Movshon, 1982). Interestingly, every plaid has an interpretation as a single translating frame, that is, as the rigid translation of a snapshot of the plaid. At sufficiently slow speeds (i.e., sufficiently low temporal frequencies of the component gratings), this rigid movement can be perceived. However, at higher temporal frequencies, the component gratings of a plaid may be perceived to move transparently in their component directions. Sometimes plaids are perceived as moving in the direction of the vector sum of the component direction-speed (velocity) vectors. And sometimes plaids are perceived as moving in yet other directions. Since Adelson and Movshon (1982) introduced plaids, there has been an enormous concern with the perceptions produced by plaids: more than 650 publications, not including abstracts, according to Google Scholar in February 2019. But there has not yet been a successful theory to explain the wide range of perceptions produced by plaids. Here, we bring forward two new assumptions and corresponding methods for the analysis and explanation of plaid motion: (a) Contrary to intuition, motion direction and mo- AQ: 9 tion speed are computed separately by the visual system—here, we concentrate exclusively on motion direction; and (b) Of the three human motion-direction computations, same-spatial-frequency plaids activate only the first- and third-order motion-perception systems.

We vary temporal frequency to selectively stimulate each one of the motion-perception systems individually and thereby to learn the rules that govern how each system processes plaids. Then, by stimulating both systems concurrently, we observe how the first- and third-order systems combine their outputs.

George Sperling, Peng Sun, Dantian Liu, and Ling Lin, Department of Cognitive Sciences, University of California, Irvine.

A poster on Experiment 2 was presented at the 2008 annual meeting of the Association for Research in Vision and Ophthalmology (Liu & Sperling, 2008). A talk on a preliminary version of contrast-strength-vector summation theory was presented at the 2009 annual meeting of the European Conference on Visual Perception (Sperling & Liu, 2009).

Correspondence concerning this article should be addressed to George Sperling, Department of Cognitive Sciences, University of California, Irvine, 3151 Social Science Plaza, Irvine, CA 92697-5100. E-mail: sperling@uci.edu

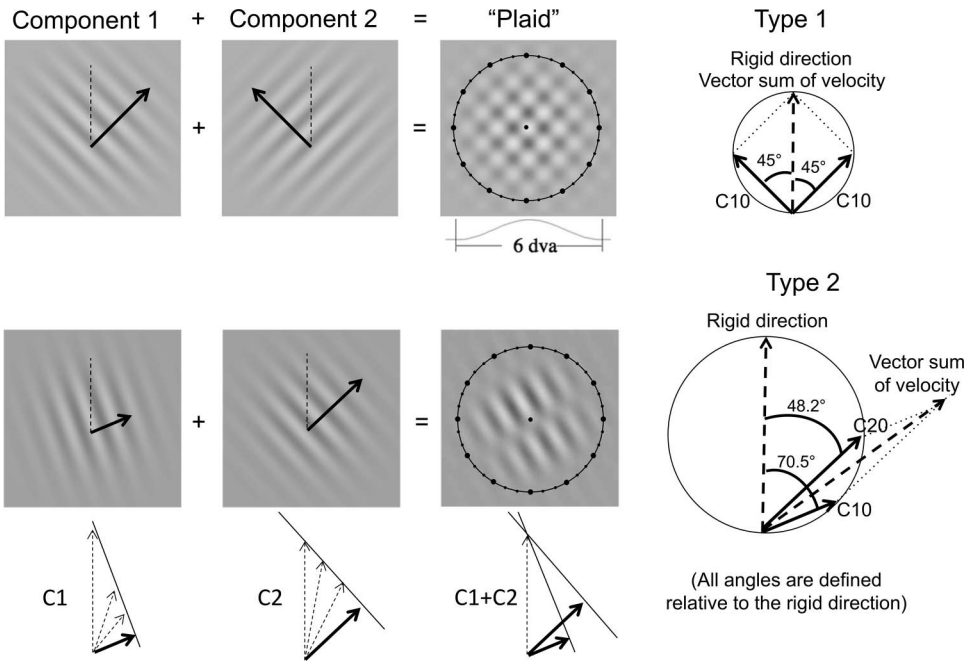
The equations governing the first- and third-order motion systems and their interaction yield a purely theoretical, parameter-free prediction of the full range of perceptions produced by same-spatial-frequency plaids. The parameter-free theoretical predictions are not perfect; without parameters, they obviously cannot describe individual differences or various subtleties in the data, but they clearly capture the essence of same-spatial-frequency plaid motion perception.

Type I and Type II Plaids

AQ: 10 Adelson and Movshon (1982) observed the motion of two kinds of plaids that Wilson, Ferrera, and Yo (1992) divided into two classes: Type I and Type II (see Figure 1). In Type I plaids, the velocity of the component gratings is such that the direction of the direction of rigid movement lies between the directions of the two component directions (see Figure 1). In Type II plaids, the direction of rigid movement lies outside of the angle formed by the two component velocity vectors (see Figure 1), making Type II plaids useful for discriminating theories. Neither the rigid direction of a plaid nor the vector sum of component velocities depend on the contrast of the component sine-wave gratings. Here, we vary the contrasts of the components of Type I and Type II plaids and use temporal frequencies above 10 Hz to exclude the third-order motion perception mechanism (Lu & Sperling, 1995a). In this restricted domain, in which only the first-order motion mechanism AQ: 11 is active, it is possible to arrive at a simple theory of the perceived direction of plaid stimuli composed of two gratings with the same spatial frequency that vary in their relative contrasts, temporal frequencies (speeds), and the angle between them.

The Aperture Problem Versus the Direction of Rigid Translation

There is intrinsic ambiguity in determining the motion direction of a one-dimensional stimulus, such as a sine-wave grating. Con-



AQ: 68 *Figure 1.* The Type I and Type II plaid stimuli and their components used in the experiments. Top row: Type I plaid. Top left: A single frame (snapshot) of Component 1, a 1 cycle per degree (cpd) sine wave with 30% contrast modulation around the mean background level that moves upward to the right within a Gaussian window (σ = 2.0° of visual angle). The temporal frequency was 10.6 Hz, resulting in a speed of (10.6 Hz)/(1 cpd) = 10.6°/s. Component 2 moves upward to the left with the same parameters as Component 1. "Plaid" is the algebraic sum of Components 1 and 2. The moving plaid appears within a circle of 6° diameter with a central fixation spot intended to control fixation, vergence, and accommodation. The ticks on the circle help the trained AQ: 69 subjects to indicate the direction of perceived movement in degrees (0°, . . . , 359°). The direction of rigid translation (also known as *rigid direction*, *direction of pattern motion*, *and intersection of constraints*) is represented by the dashed arrow in the Type I diagram. The rigid direction was randomly varied from trial to trial between 0°, . . . , 359°. The labels C10 and C20 indicate components with temporal frequencies of 10.6 Hz and 21.2 Hz, respectively. Middle row: Type II plaid. The components' arrows represent velocity (direction and speed). Bottom row (C1, C2): The vectors indicate the range of velocities consistent with each of the Type II components (C1 + C2). The geometric construction of possible velocities consistent with each component (the intersection of constraints) shows that one, and only one, pair of component velocities is consistent with a rigid translation of the plaid pattern. See text for details and see the online supplemental materials for video examples of Type I and Type II plaids.

sider a snapshot of a sine-wave grating displayed on a piece of paper, and the paper is set into motion. Observing through a circular window, the motion is perceived as being perpendicular to the orientation of the grating no matter what arbitrary direction the piece of paper may be physically moving in. Indeed, all directions of motion of the paper that happen to have the same motion component perpendicular to the stripes of the grating produce precisely the same image inside the aperture, as illustrated by the motion vectors in Figure 2a. This is the "aperture problem."

If a second grating with a different orientation is added to the first grating, forming the pattern known as a *plaid* (Adelson & Movshon, 1982; Movshon, Adelson, Gizzi, & Newsome, 1985), the ambiguity is resolved. Any plaid—independent of the spatial and temporal frequencies of the component sine waves—can be equally well represented by a snapshot of the plaid on a piece of paper and moving the paper in a unique direction. We call the motion direction of the moving paper that reproduces the component sine waves the *rigid direction* for the obvious reason that it represents the plaid motion as the translation of a rigid object, that is, the snapshot of the plaid. The direction of rigid translation is also called the direction of *pattern motion*. The direction of rigid translation, the rigid direction, is a purely physical concept. When we say subjects perceive motion in the rigid direction, that refers only to the judged direction. It has absolutely no implication about perceived rigidity of the moving stimulus.

When viewing a plaid (e.g., Figure 2b) through an aperture, in addition to an algebraic solution, there is a simple geometric construction, the intersection of constraints (IOC), for finding the direction of rigid motion (Adelson & Movshon, 1982). The IOC construction is illustrated at the bottom of Figure 1.

Basic Motion Perception Systems

A basic visual motion perception system takes as its input a scalar property of dynamic visual stimulus. This scalar property, for example, luminance, is a function of x, y, t. The motion system produces as its output a time-varying vector flow field, that is, a vector with a contrast strength and direction that is also defined at each point of x, y, t. That is, a motion system converts an x, y, t cube of scalars into an x, y, t cube of vectors. The cortical area MT is an example of a representation in the brain of a motion vector flow field.

Three motion systems have been proposed for human vision (e.g., Lu & Sperling, 1995a). They differ primarily in their inputs to the motion computation versus differing in the motion computation itself, which may be quite similar in different motion-perception systems. The first-order system (often misleadingly called the *luminance motion system*) takes as its input not luminance as a function of x, y, t (luminance is always a positive quantity) but rather the Weber contrast of each point in x, y, t, a quantity that is positive for points more intense than their spatial surround and negative for points dimmer than their surround (Reichardt, 1961). The second-order system takes as its input the local texture contrast in the neighborhood of each point x, y, t (i.e., the variance of luminance or, equivalently, variance of point contrast). The third-order system takes as its input the salience value at each point x, y, t. Salience is a complex computation that, like figure-ground, is influenced by attention. Salience is large for points

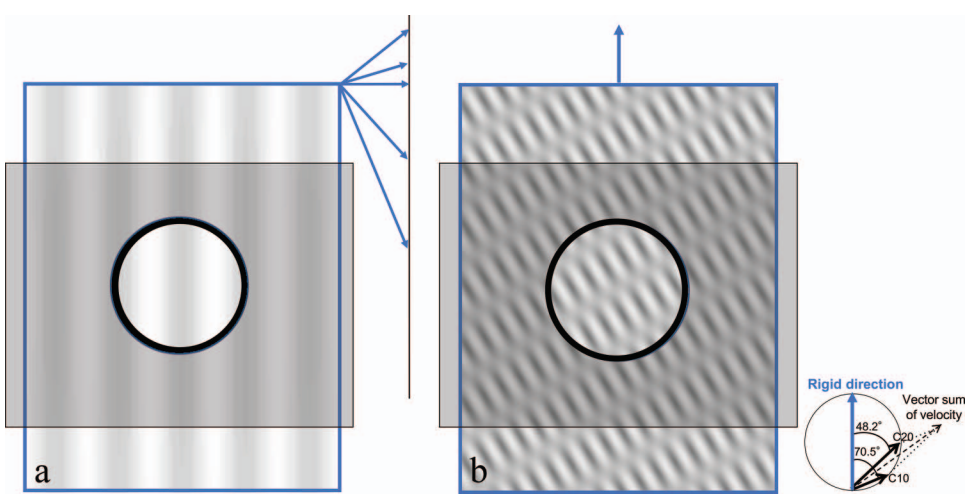


Figure 2. Rigid motion direction defined: All motions of sine-wave gratings and of plaids (pairs of sine-wave gratings) can be produced by viewing a moving snapshot of the grating or plaid through an aperture. (a) A sine-wave grating. The arrows indicate velocities (directions and speeds) of motions of the rectangular picture that would produce identical image sequences within the circular aperture. Therefore, the physical direction of motion of a sine-wave grating is inherently ambiguous. (b) The Type II plaid used in the experiments. When the two moving sine-wave grating components of a plaid are nonparallel, there is a unique direction and velocity of the snapshot of the plaid (the direction of rigid translation) that, within the window, reproduces exactly the two different velocities of the component gratings. The black arrows in the insert show the velocities of the plaid component sine waves; the dotted arrows show the rigid direction and the vector sum of velocities direction. Although a brief view of a moving plaid is logically sufficient to define the rigid direction, that is, the direction in which the snapshot of the plaid is moving, the rigid direction usually is not the perceived direction of motion. See the online article for the color version of this figure.

that are perceived as figure and small for points interpreted as background.

Factors That Determine Perceived Motion of Plaids

There have been numerous investigations of the factors that determine the perception of plaid motion, including contrast, spatial and temporal frequencies, viewing duration, and other factors (e.g., Bowns, 1996, 2018; Stone, Watson, & Mulligan, 1990; Weiss, Simoncelli, & Adelson, 2002; Wilson et al., 1992). Several alternative computational theories have been proposed. In the vector summation of velocities algorithm, the physical velocities of the components, or their perceived velocities, are first computed individually and then combined by vector summation (Champion, Hammett, & Thompson, 2007; Stone et al., 1990; Wilson et al., 1992; Wilson & Kim, 1994; Yo & Wilson, 1992) to produce the perceived output. AQ: 13

A feature tracking explanation of plaid motion perception was proposed initially by Adelson and Movshon (1982) and more recently by Bowns (2018) and others. Note that feature tracking is an instruction to the subject, not a computational motion mechanism. Attention to a feature makes it more salient (Blaser, Sperling, & Lu, 1999; Lu & Sperling, 1995b; Tseng, Gobell, & Sperling, 2004); the third-order motion system computes motion on a spatiotemporal salience field in the same way that the first-order system computes motion on a spatiotemporal point-contrast field. That is, except for the input and some parameters, the basic motion computation is the same for the first- and third-order systems. Without a motion signal to indicate which way the to-be-tracked feature is moving, tracking a feature would require a feature search to find the to-be-tracked feature every time it moved. For present purposes, the most significant property of the third-order motion system is that its sensitivity declines rapidly at temporal frequencies above 3 to 4 Hz. For most subjects, third-order motion perception is insignificant above 10 Hz (Lu & Sperling, 1995a).

On the other hand, sensitivity of first- and second-order motion is preserved for frequencies up to about 10 Hz and only declines as frequencies are increased above 10 Hz (Lu & Sperling, 1995a). In the present study, we wish to study plaids that stimulate only the first-order (and possibly the second-order) system and to bypass the third-order motion system. Therefore, we use stimuli with temporal frequencies of 10 and 20 Hz in a wide range of contrasts. We will demonstrate that high-temporal-frequency plaids yield a relatively simple and consistent theory of (first-order) plaid motion. Subsequently, to measure the influence of third-order order motion system, stimuli with temporal frequencies as low as 1 Hz will be used. AQ: 14

Although there is clear evidence that component contrasts play an important role in plaid motion perception (Champion et al., 2007; Stone et al., 1990), it is not clear how different contrast ratios between the two components, nor how overall contrast levels, determine the perceived direction of plaid motion. Here, we systematically study the effect of contrast on the perceived direction of plaid motion.

Present Study

AQ: 15, 16 Four experiments were conducted to investigate the perceived motion direction of plaid stimuli that have sufficiently high temporal frequencies such that they were expected to stimulate only the first-order motion-perception system. All plaid sine-wave components are 1 cycle per degree (cpd). Experiment 1 compared motion-direction judgments in two paradigms to determine which was more appropriate to use for Experiments 2 to 4. Experiment 2, the main experiment, investigated the perceived direction of Type II plaid stimuli having a 10-Hz and a 20-Hz component that vary over the full range of joint contrasts, plus two representative low-temporal-frequency stimuli. The results of Experiment 2 were incorporated in a contrast-strength-vector summation theory that was used to predict the perceived direction of new plaid stimuli. The model was tested in Experiment 3, which investigated perceived direction of plaids composed of the same 10-Hz and 20-Hz and 1 cpd frequency components as in Experiment 2, except that the angle between the two components was varied over the full range. Experiment 4 was a control experiment to determine the extent to which motion-direction judgments could have been determined merely from the perceived spatial orientation of the plaid, that is, from a judgment that can be made in a static display. Finally, a pure theory, with zero parameters estimated from the data, was shown to capture the essential features of the perceived direction of same-spatial frequency plaids.

General Method

Direct Estimation of Motion Direction

Perhaps the most sensitive way of measuring the perceived direction of motion is to present stimuli that vary slightly in motion direction around a fixed direction, for example, vertical. The subject's task would be to report whether the direction of a given stimulus is to the left or right of vertical. We investigate a minor variant of this procedure, but we also designed a direct reporting method for measuring perceived motion direction in any direction from 0° to 359°. Subjects were trained, with feedback, to directly estimate motion directions of moving sine-wave gratings in degrees (with the aid of tick marks in a circle around the stimulus) and to type their estimates on a keyboard. In principle, direct estimation is similar, for example, to the method used by Cropper and Badcock (2008) in which subjects used a computer mouse to indicate motion direction. Our subjects quickly learned to directly estimate motion directions in degrees. Training with the training stimuli continued until their judgments were quick and accurate. Whereas the training trails had error-correction feedback, the experimental trails did not have correct answers and, therefore, there was no feedback (Sperling, 1992; Sperling, Dosher, & Landy, 1990). Before each new block of trials, subjects were shown sample stimuli to familiarize themselves with the stimuli they were to judge in that session. AQ: 17

Initially, subjects were given various optional responses to indicate that they were consciously combining two perceived directions of motion or were perceiving ambiguous motion. As there was never such a report, which also coincides with the experimenters' observations, these additional response options were discontinued.

Stimuli

All stimuli in these experiments were composed of sine-wave gratings (sinusoidal modulations around the mean luminance); two

such gratings were added together to form a plaid (Figures 1 and 2). The spatial frequency of all gratings was 1 cpd of visual angle. In most conditions, the temporal frequencies of the gratings were 10 Hz or greater. The third-order motion system begins to lose sensitivity at about 4 Hz and is relatively quite weak at frequencies of 10 Hz or greater. Therefore, we expect these stimuli to minimize the contribution of third-order motion system to the perception of motion direction and, thereby, to stimulate only the first-order motion system or possibly the first- and second-order systems. The success of this manipulation will be evident from the data.

All stimuli were viewed within a Gaussian window with a standard deviation of 2 degrees of visual angle (dva), which faded out in 4 to 5 dva, as illustrated in Figure 1. The entire windowed stimulus appeared within a black circle, serving the purpose of facilitating vergence. The circle subtended 6 dva. The peripheral markers on the inside edge of the vergence circle are references for subjects to make motion direction estimates in degrees, from 0° to 359°. A central fixation spot and the vergence circle with its markers were always present. Subjects maintained fixation at the central spot throughout a trial.

Experiment 1: Comparison of Two Paradigms for Estimating Motion Direction

Method

AQ: 19 Using our method of direct direction estimation, we compared two paradigms: (a) *random directions*, in which the direction of rigid motion on each trial is chosen randomly from the whole circle (0°, 359°); and (b) *restricted directions*, in which the direction of rigid motion is chosen mainly from close-to-vertical directions (−4°, 4°). Random direction procedures are typically used in more cognitive tasks (e.g., Emrich, Riggall, LaRocque, & Postle, 2013), whereas restricted directions are typically used in more psychophysical tasks (e.g., Cropper & Badcock, 2008; 2AFC plaid motion task). Although random directions have been used in human motion tasks (e.g., van Maanen et al., 2012), the two psychophysical methods have not been compared to determine their differential suitability for judging motion direction (our task).

Stimuli. Symmetrical Type I plaids were used with component spatial and temporal frequency at 1.0 cpd and 10.6 Hz, respectively. The two components were oriented at 90° relative to each other. In order to minimize third-order motion perception, we used a high temporal frequency, 10.6 Hz, and low-contrast component gratings. Because the human third-order motion system has greatly reduced sensitivity at 10 Hz (Lu & Sperling, 1995a), and also at low contrasts (i.e., below about 5%), one plaid component, the higher contrast component, always had a contrast of 2%. The other component had contrasts of 2%, 1%, 0.5%, 0.25%, and 0%. The 90° angle between gratings causes the second-order motion system to be ambiguous (see the "Second-order motion contributions to plaid motion" section), so the second-order motion system is useless for these stimuli. These stimuli are designed to and expected to excite only the first-order motion system.

A total of 10 plaids were created as follows: Five combinations of contrast pairs were used (i.e., [2%, 2%], [2%, 1%], [2%, 0.5%], [2%, 0.25%], [2%, 0%]). Each contrast pair was used to create two mirror symmetric plaids. For pairs of equal contrast (i.e., 2%, 2%), the same plaid was used twice. Figure 3 shows snapshots of the F3 plaids in one of their possible orientations (rigid direction vertical).

Figure 4 (top row) shows vector diagrams of the two component F4 gratings and of the resulting plaid as they occur in the two parts of the experiment. The rigid direction and the vector sum of the component velocities both point exactly between the vectors representing the two component gratings. Both the rigid direction and the vector sum of velocities are independent of the contrasts of the component gratings as long as the contrasts are not zero.

Apparatus. Stimuli were generated on an Apple Mac G5 computer using a MATLAB 7.04 (Mathworks, Inc.) program with Psychtoolbox (Brainard, 1997; Pelli, 1997). Stimuli were displayed on a Hyundai RBG monitor with 1024 × 768 resolution. The frame rate was 85 Hz. The luminance of each pixel was resolved with 10-bit accuracy. A standard lookup table was generated by means of a psychophysical procedure that linearly di-

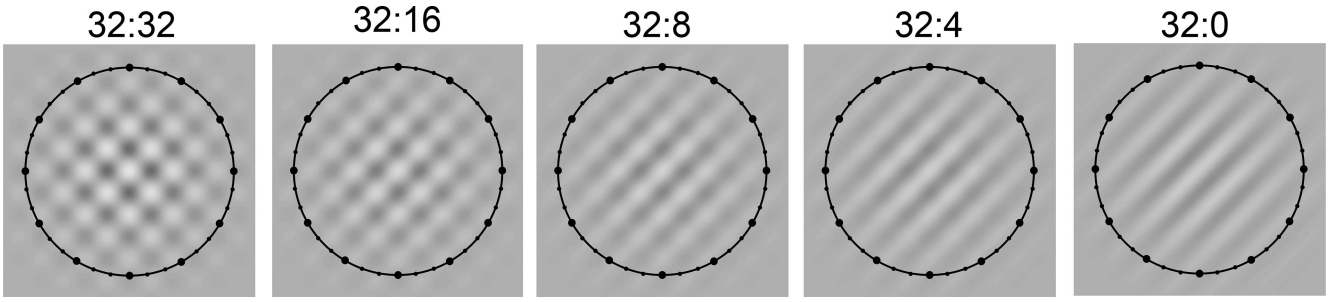


Figure 3. Snapshots of five of the nine Type I plaid stimuli used in Experiment 1. The other four are mirror images of the asymmetric stimuli. The two sine-wave components move with equal velocities in the directions +45° and −45°. The contrast amplitudes of the illustrated components are indicated as percents of the maximum achievable contrast. For the illustrated plaid orientation, and for all indicated contrast combinations except 32:0, both the direction of rigid movement and the vector sum of velocities of the sine-wave components are vertical (as in Figure 1). The stimulus 32:0 is a single grating, not a plaid; the rigid direction is undefined (Figure 2a). For illustrative purposes, the stimulus contrasts shown here are relative to a maximum contrast of 32%. In Experiment 1, the same contrast ratios were used, but the maximum contrast was 2%. See text for details and see the online supplemental materials for video examples of Type I and Type II plaids.

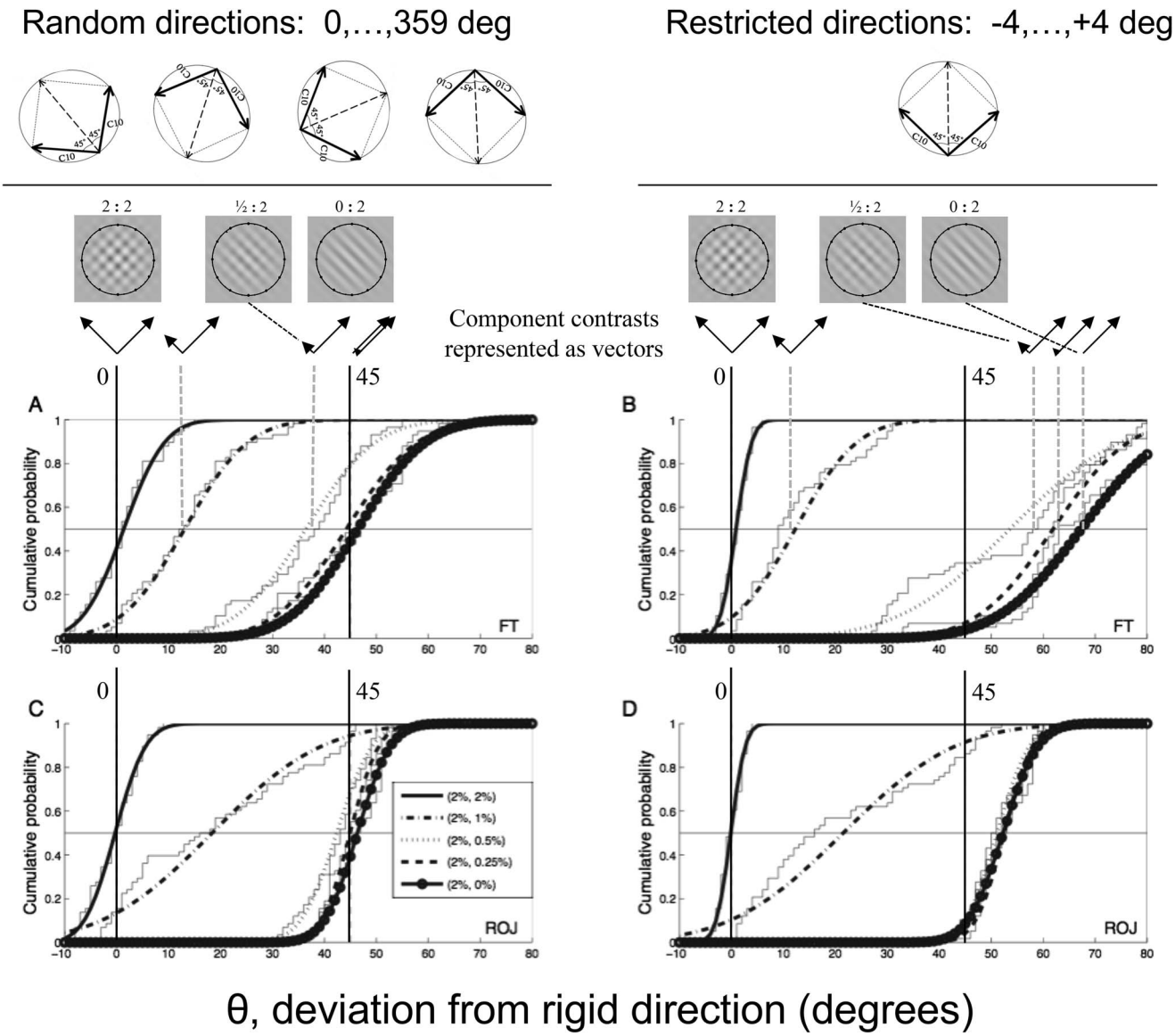


Figure 4. Experiment 1. Cumulative histograms of the judged directions of plaids in which the direction of rigid motion varies randomly between (0°, 359°) or between (−4°, +4°). Thumbnail images represent three of the five combinations of component contrasts. The illustrated contrasts, about 30%, are much higher than the presented contrasts (max = 2%). Below the thumbnail stimulus images, the contrasts of the components are represented as vectors: Vector direction represents component direction; vector length represents component contrast. The numbers 0 and 45 and the corresponding vertical lines represent the only two stimuli whose physical directions and judged directions are expected to coincide within measurement error, as they indeed do in the random motion-directions (0°, 359°) but not in restricted motion-directions (−4°, +4°) paradigm. (A, B) Data from subject FT. (C, D) Data from subject ROJ. In each panel, the abscissa indicates the direction of motion θ relative to the rigid direction; the ordinate indicates the cumulative probability that a judged direction is greater than θ. Jagged curves indicate raw data; smooth curves indicate Gaussian fits. The five curves—some overlap completely—represent the five contrast ratios of the stimulus components. All directions are relative to the rigid direction, which is represented here as upward vertical. See text for details. Single-component plaid directions (pure sine waves) are judged correctly for the random motion-directions (0°, 359°) conditions (A and C) but incorrectly for the restricted motion-directions (−4°, +4°), conditions B and D. The two lowest contrast components have a very small influence on perceived direction.

vided the whole luminance range into 256 gray levels (Lu & Sperling, 2001). The mean luminance of the monitor was 118.6 cd/m^2. Subjects viewed the display binocularly at a distance of 50 cm in a darkened room.

Subjects. Two subjects aged 20 and 23 years of age, with normal or corrected-to-normal vision, participated in this experiment. Subjects gave informed consent to the experimental procedure, which was approved by the Institutional Review Board of the

University of California, Irvine, prior to their participation. Subjects were paid for participation.

Procedure. Subjects initiated an experimental session by pressing a button on the keyboard. A plaid stimulus was then displayed on the monitor for 200 ms, followed by a screen containing only the vergence circle, the peripheral marker, and the central fixation spot. Subjects (previously trained) were instructed to estimate the overall plaid motion direction in an integer number of degrees and to type their estimate on the computer keyboard. Unbeknownst to the subjects, pure sine-wave stimuli that were interleaved throughout the experiments were similar to their prior training stimuli.

Sessions lasted approximately 1 hr. Two experimental conditions were conducted separately in each experimental session: random motion-directions (0°, 359°) and restricted motion-directions (−4°, +4°). In random motion-directions (0°, 359°), the whole 360° circle was equally divided into 18 sectors, with each sector spanning 20°. Then, a direction (in integer degrees) was randomly chosen from a uniform distribution within each 20° sector. In restricted motion-directions (−4°, +4°), each of the nine close-to-vertical directions of motion (−4°, −3°, −2°, −1°, 0°, 1°, 2°, 3°, 4°) was used twice. A total of 36 directions (18 for the random-directions condition, 18 for the restricted-direction condition) was generated for each session. In an experimental session, each of the 10 plaids of the current paradigm (ones in Figure 3 and their mirrored images) was presented in each of the 36 rigid directions in a pseudorandom order. For each condition (random, restricted), 360 trials split into two sessions were randomly presented to each subject.

Data analysis. All data were standardized to have the rigid direction as 0° by subtracting the rigid direction from each observation. Given the symmetrical nature of Type I plaids, results from different-contrast mirror-image plaids were pooled after flipping one of the plaids around the rigid direction (0°). Results for plaids with equal contrast components were simply pooled together. Thereafter, the upper and lower 10% of the data were excluded from further analysis as outliers. The remaining 80% data were fit (minimize $\|p - \hat{p}\|^2$) with a cumulative Gaussian. The results are shown with the cumulative Gaussian curve overlying the empirical probability distribution plot. The mean and standard deviation of the fitted Gaussian distribution were used to summarize the data.

Results

Psychometric functions. The left and right halves of Figure 4 show the data obtained in the two different range-of-motion-direction protocols, (0°, 359°) and (−4°, +4°), as indicated by the vector diagrams in the top row. Actual plaid directions are normalized relative to the rigid direction, and mirror-image plaid motions are flipped so that the motion direction of the lower contrast component is always −45° and that of the higher contrast component is +45°.

The five curves in Figure 4 represent individual plaid contrast ratios illustrated in Figure 3. It is immediately evident that in both protocols, both subjects judge equal contrast plaids (2%, 2%) as moving in the rigid direction (0°), which, for these plaids, is also the direction of the vector summation of velocities. However, plaids with a contrast ratio of 2:1 (2%, 1%) are perceived as moving in a quite different direction (neither rigid nor vector sum of velocities) but a direction closer to the direction of the higher contrast component. Plaids with component contrast ratios of 4:1, 8:1, and a pure sine wave are all judged as moving in virtually the same direction, that is, the direction of the higher contrast component totally dominates. For these already low-contrast plaids, the two lowest contrast components have very little effect.

Context effects on motion-direction judgments. Single-component directions (2%, 0%) are judged correctly for (0°, 359°) conditions (Figure 4A, C) but are judged incorrectly for the restricted motion-directions (−4°, +4°) conditions (Figures 4B, D). In particular, for the motion-directions (0°, 359°) condition, the judged single-component direction was 46.0° for subject FT (Figure 4A) and 45.5° for subject ROJ (Figure 4B); both judgments are within measurement error of 45°. However, for the (4°, +4°) condition, large deviations from the veridical 45° direction were observed: 64.2° for subject FT (Figure 4B) and 52.3° for subject ROJ (Figure 4D). This means that, in the context of many nearly vertically moving plaids, the appearance of deviations from the vertical direction is exaggerated.

Discussion

Context versus adaptation. The consistent overestimation of deviations from the most prevalent motion direction in the restricted-directions paradigm means that the context of recently perceived motion directions greatly influences the perception of the current stimulus. This is analogous to a motion-adaptation paradigm in which exposure to a particular orientation biases subsequent tests away from the adapted orientation (e.g., Barlow & Foldiak, 1989). However, the procedure involves only very brief trials of very low-contrast stimuli separated by longer intervals during which responses are recorded. This is quite different from the usual adaptation procedures in which high-contrast stimuli are presented continuously for long periods. The context dependence observed in the restricted-directions paradigm seems to involve higher level processes than classical orientation adaptation. It is consistent, for example, with the notion that the visual system adapts to maximal sensitivity, in this case, to best discriminate among the most frequent inputs.

The context dependency reported here is different from the systematic bias in orientation estimation tasks, in which subjects' orientation estimations are biased either away from cardinal orientations (repulsion) or toward cardinal orientations (attraction) depending on noise type and its magnitude (Tomassini, Morgan, & Solomon, 2010; Wei & Stocker, 2015). In those tasks, systematic biases are never found for oblique orientations ±45°; the Bayesian explanations of those biases also predict zero bias for ±45° orientations. The very large contextual bias in ±45° orientation judgments produced by our restricted-directions paradigm appears to be unique and worthy of further study. For the purpose of this study, however, the bias in the restricted-directions paradigm forces the use of the (0°, 359°) random-directions paradigm.

Context adaptation probably occurs in both paradigms. In the restricted-directions paradigm, the vertical context exerts a specific directional bias, the magnitude of which depends on the subject. In the (0°, 359°) random-directions paradigm, there is no systematic bias because the preceding trials are randomly oriented relative to the current trial. Systematic sequential analysis (of a much larger

data set) should reveal a small build-up of random bias even in the (0°, 359°) random-directions paradigm.

The large effect of contrast on perceived direction. Van Santen and Sperling (1984) showed that the elaborated Reichardt model accounted astonishingly well for six counterintuitive predictions of phenomena of first-order sine-wave motion. The motion energy model (Adelson & Bergen, 1985) and the Hilbert transform model (Watson & Ahumada, 1985) were shown to be functionally equivalent to the Reichardt model (van Santen & Sperling, 1985). The output of all these equivalent detectors is proportional to the square of the amplitude (contrast) of an input sine wave. Speed is important for a Reichardt detector only insofar as it produces a temporal frequency that is effective for that particular detector. Greater or smaller speeds than the optimum produce smaller outputs. As the plaids in Experiment 1 are composed of independent high-temporal-frequency components that presumably are detected by Reichardt detectors, and as temporal frequency was not varied, contrast was expected to be, and indeed was, the sole determiner of perceived direction.

Practical Conclusions

In Experiment 1, all temporal frequencies were above 10 Hz and the contrasts of all plaid components were less than or equal 2%. Even though the contrast was too low and movement too fast to clearly perceive moving features, subjects were perfectly able to judge the motion direction of these stimuli. Whereas a restricted-directions paradigm enables very precise ordinal comparisons between motion directions, the (0°, 359°) random-directions paradigm gives a context-neutral measurement of the perceived motion direction of plaid stimuli and enables metric comparisons between perceived motion directions. Our subsequent experiments use the (0°, 359°) random-directions paradigm. Before considering the theoretical implications, more data are needed.

Experiment 2: Perceived Motion Direction of Type II Plaids as Function of Component Contrast and Temporal Frequency

Method

Experiment 2 is an extensive empirical study of mostly high-temporal-frequency Type II plaids that are designed to stimulate only the first-order motion system with temporal frequencies ranging from 10 to 30 Hz. It explores the full range of absolute and relative contrasts of the components. Following the pure first-order phase of the experiment, to illustrate the influence of other motion perception systems, a limited number of plaids with temporal frequencies as low as 1.0 Hz are tested.

Figure 1 shows snapshots of a Type I plaid stimulus from Experiment 1 and a Type II plaid stimulus from Experiment 2. The diagrams in Figure 1 make obvious the advantages, originally noted by Ferrera and Wilson (1990), of Type II plaids for discriminating between motion theories, particularly between theories that predict perceived direction in direction of rigid translation versus predicted perceived direction equal to the vector sum of velocities. For a symmetrical Type I plaid, as in Figure 1, the directions of rigid translation and of the vector sum of component velocities are identical. For the Type II plaid, the directions of rigid translation and vector sum of velocities are very different.

Stimuli. Type II plaids were composed of two components, both sine-wave gratings, designated here as C10 and C20. In the main portion of the experiment, these components moved with temporal frequencies 10.6 Hz and 21.2 Hz. C20 moved 48.2° clockwise relative to the rigid direction; C10 moved 70.5° clockwise relative to the rigid direction (see Figure 1). Both components had a spatial frequency of 1.0 cpd. C20 translated 90° in phase between consecutive frames; C10 translated 45° in phase between consecutive frames. Stimulus duration was 200 ms, that is, a total of 17 stimulus frames at a frame rate of 85 Hz.

A total of nine different component contrast ratios (Figure 5a) F5
was explored times four conditions with different maximum contrasts. Contrast ratio is the contrast of the lower temporal fre- AQ: 24
quency component divided by the contrast of the higher temporal frequency component, which is simply Contrast(C10)/Contrast(C20). The contrast ratio took values of 0, 1/8, 1/4, 1/2, 1/1/, AQ: 25
2, 4, 8, and infinity (i.e., 8/0). Within a condition, the higher contrast component in every ratio had the maximum contrast value that defined the condition. For the four conditions, the maximum values were either 32%, 16%, 8%, or 4%. Figure 5a shows the full range of stimuli for the 32% contrast condition. Figure 5b shows the cumulative response histograms for leftmost five stimuli of Figure 5a.

Procedure. The data were collected in two phases, which differed slightly in the way in which pseudorandom directions were chosen. In Phase 1, a circle was equally divided into 36 intervals of 10. A direction (in integer degrees) was randomly chosen from a uniform distribution within each 10° interval, thereby producing 36 different directions. In Phase 2, a direction was chosen from each 4° interval around the circle, yielding 90 different directions. Trials were run in a pseudorandom mixed-list design; that is, from a subject's point of view, on any trial, any one of the tested stimuli was as likely to be presented as any other. Different combinations of temporal frequencies, however, were tested in different sessions. Each experimental condition (contrast of C10, contrast of C20) contained one trial from every direction; data from the two phases were combined, yielding 126 trials per data point.

For one subject, the entire main experiment was repeated at a frame rate of 120 Hz, and with component temporal frequencies of 30 Hz and 15 Hz replacing the original 21.2 Hz and 10.6 Hz.

Subsequent to the main experiment, two separate sessions were run that used only the highest contrast stimuli, in which one component's contrast was always 32%: The temporal frequencies of the plaid components were (a) 10.6 Hz and 5.3 Hz (designated as 10:5 Hz) with a stimulus duration of 765 ms, and (b) 2.02 Hz and 1.01 Hz (2:1 Hz) with a stimulus duration of 1,000 ms.

The point of subjective equality (PSE) is the median of the probability density function of the judged directions. Because outlying data are likely to come from causes not of interest (typing errors, momentary inattention, etc.), to obtain a more robust measure of central tendency, the upper and lower 10% were excluded from analysis (to remove outliers and obtain a more robust measure of central tendency). The mean (equal to the median) was estimated from a cumulative Gaussian curve fitted to the data.

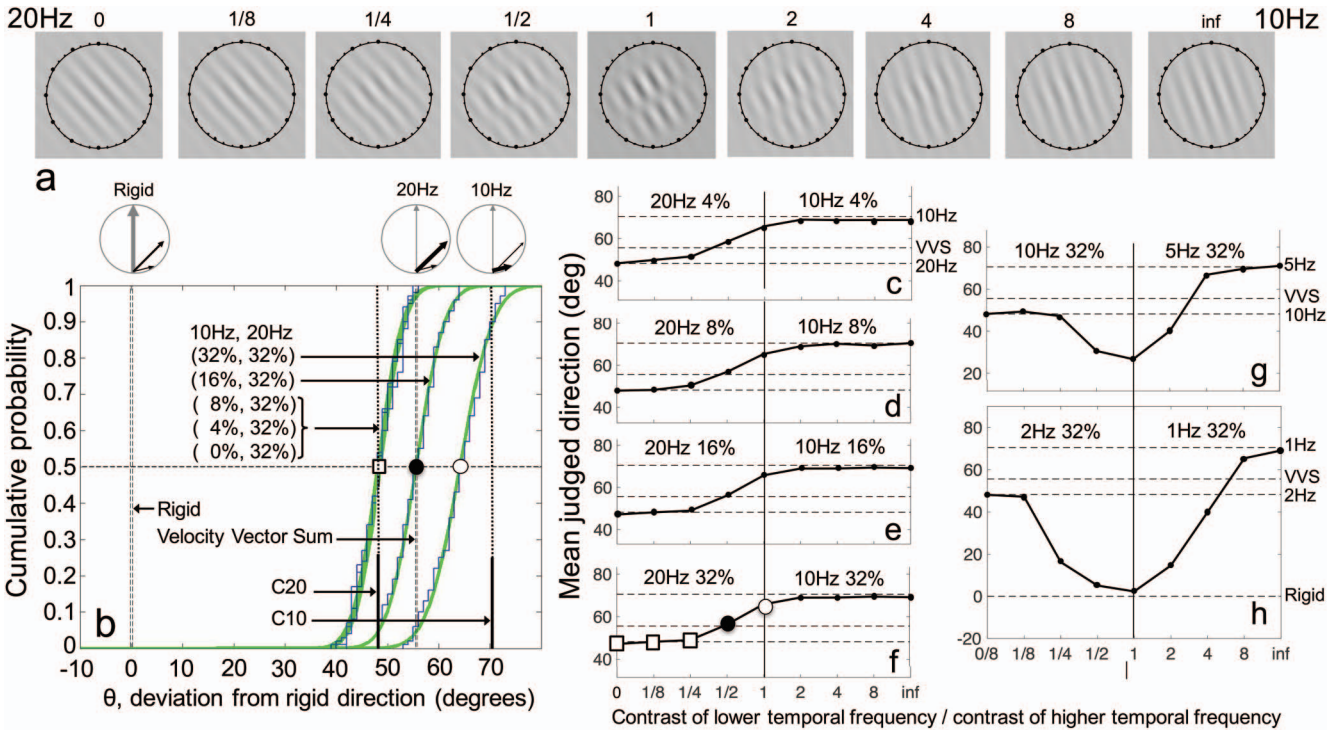


Figure 5. Stimuli and results for one subject in Experiment 2: Judged direction of randomly oriented Type II plaids as a function of the contrasts and temporal frequencies of the components. (a) The 32% contrast set of nine stimuli and the contrast ratios: (lower temporal frequency)/(higher temporal frequency). (b) Cumulative histograms of direction judgments for plaids with five different component contrast ratios. Abscissa: Judged direction θ relative to the rigid direction. The angles of the two component gratings (48.2°, 21.2 Hz; 70.5°, 10.6 Hz) and the rigid direction (0.0°) are represented as dark arrows in the diagrammatic thumbnails above vertical dotted lines that extend to the abscissa. Ordinate: Cumulative probability that the judged direction of motion is greater than θ. Three cumulative Gaussian curves overlie the five data sets. Data for plaids with contrast ratios 0/32, 4/32, and 8/32 are all equivalent and are all fit by the same Gaussian curve. The point of subject equality (PSE) is taken as the angle where the cumulative probability is 0.5, which, for the smallest three contrast ratios, corresponds almost exactly to the direction of the 20.2-Hz grating (indicated as C20). The PSEs are indicated by different symbols, and are again represented in Panel f. (c–h) Each panel exhibits motion direction judgments for nine plaids of different contrast ratios. All plaids in a panel have the same pair of component temporal frequencies and same maximum contrast. The temporal frequency of the component that has the maximum contrast in each half of the graph is indicated. Velocity Vector Sum (VS) is the sum of the plaid component velocities; Rigid is the direction of rigid translation; both are independent of the contrast ratio. Whereas the temporal frequencies in Panels c through f are 10 Hz and 20 Hz (pure first-order stimuli), the plaid with equal-contrast components of 1 Hz and 2 Hz (panel h) stimulates pure third-order motion precisely in the rigid direction. See the online supplemental materials for video examples of the above plaids. See the online article for the color version of this figure.

Results and Discussion

Deriving PSEs. Figure 5b shows cumulative probability distribution functions (as in Figure 4) of motion direction responses for subject ROJ viewing Type II plaids composed of 1 cpd gratings with temporal frequencies of 10.6 Hz and 22.2 Hz at five contrast ratios. For the three highest contrast ratios, the PSEs are 47.1°, 48.2°, and 49.0°. These PSE are within a degree of 48.2°, the direction of C20, and are so close together that the two highest contrast ratio plaid judgments are not statistically different from the pure C20 sine-wave grating. Alternatively stated, for plaids with component contrasts of (32%,8%), and of (32%,4%), the lower contrast component has no measurable effect on motion-direction judgments.

For a contrast ratio of 2:1, the PSE occurs at 56.2°, which, at this scale, is hard to distinguish from 55.6°, the vector sum (VS) of the component velocities. However, when both gratings are of equal contrast (32%, 32%), the PSE is 64.7°. Equal-component-contrast plaid motion is judged much closer to the 10-Hz component direction than would be expected from VS.

PSEs as a function of contrast. In Panels b to h of Figure 5, only the PSEs are represented, not the probability distributions from which they were derived. The PSEs in Figure 5a reappear in the left half of Figure 5f. The standard error of the PSEs in Figure 5 averages less than 0.5° for all three subjects, that is, the 95% confidence interval is less than ±1° for these data.

As a function of the contrast ratio, the motion direction judgments shift smoothly from one component direction to the other. Neither the vector sum of the component velocities nor the rigid direction changes as a function of contrast, and they are represented as horizontal lines. When the temporal frequencies of the components are reduced to 5 Hz and 10 Hz, and the display duration increased to 765 ms, the judgments of approximately equal-contrast component plaids approach the rigid direction. At component frequencies of 1 Hz and 2 Hz, subject ROJ judges the (32%, 32%) plaid as moving precisely in the rigid direction.

The four sets of motion direction judgments in Figures 5c to 5f are remarkably similar. To better display the relations between the six data curves in Figure 5c to 5h, they are all plotted in the same graph in Figure 6a. Figure 6 also shows data from two other subjects. An astounding result in Figure 6 is that when the component temporal frequencies are 10 Hz and 20 Hz for subjects ROJ and AL, and 15 Hz and 30 Hz for subject FT, the judged direction of plaids that have the same contrast ratio is the same, independent of the absolute contrast. For example, for a contrast ratio of 1/2, the judged directions of the four plaids composed of (32%, 16%), (16%, 8%), (8%, 4%), and (4%, 2%) are statistically identical. That judged direction depends only on contrast ratio (not on the actual contrasts) obtained for all other contrast ratios, from zero to infinity.

A second significant property of the data in Figure 5b and 5d is that all the perceived directions of the 20:10-Hz plaids (two subjects) and 30:15-Hz plaids (one subject) fall between the two perceived directions of the 20-Hz and 10-Hz components individually. The contrast ratio, and only the contrast ratio, determines how much each component contributes to the perceived direction. For high-temporal-frequency stimuli (i.e., pure first-order), there is no hint of any tendency toward the rigid direction and no trace of an IOC computation.

The 20:10-Hz data for subject FT in Figure 5c differ from those for the other two subjects, in that perceived direction depended both on the contrast ratio and on the absolute value of contrast. The higher the absolute contrasts of the plaids, the more they deviate toward the rigid direction. Perceived directions of plaids with contrasts of 32% and 16% deviate so much toward the rigid direction that the perceived direction no longer lies between the directions of the components.

A subject with a more sensitive third-order motion system. We interpret the data of subject FT in Figure 6c as indicating that even at temporal frequencies of 20:10 Hz, there was a significant contribution of a third-order motion computation. Therefore, a complete additional set of trials was conducted, identical to the 20-Hz and 10-Hz trials, except that the plaid component frequencies were now increased to 30 Hz and 15 Hz. The 30:15-Hz data of subject FT exhibit the same astounding properties as the 20:10-Hz data of the other two subjects: Perceived plaid direction depends only on contrast ratio (not on absolute contrast), and all perceived directions lie between the directions of the two components. The data of the lowest contrast (4%) 20:10 plaids overlie the four 30:15-Hz data curves. The interpretation of these results is that temporal frequencies of 30:15 Hz are sufficiently high to eliminate third-order motion processing at all contrast levels. For this subject, third-order motion processing persists at 20:10 Hz until the component contrasts fall at or below 4%. On the other hand, when the component temporal frequencies are reduced to 2:1 Hz, subject FT perceives plaids with contrast ratios of 1/2 and 1

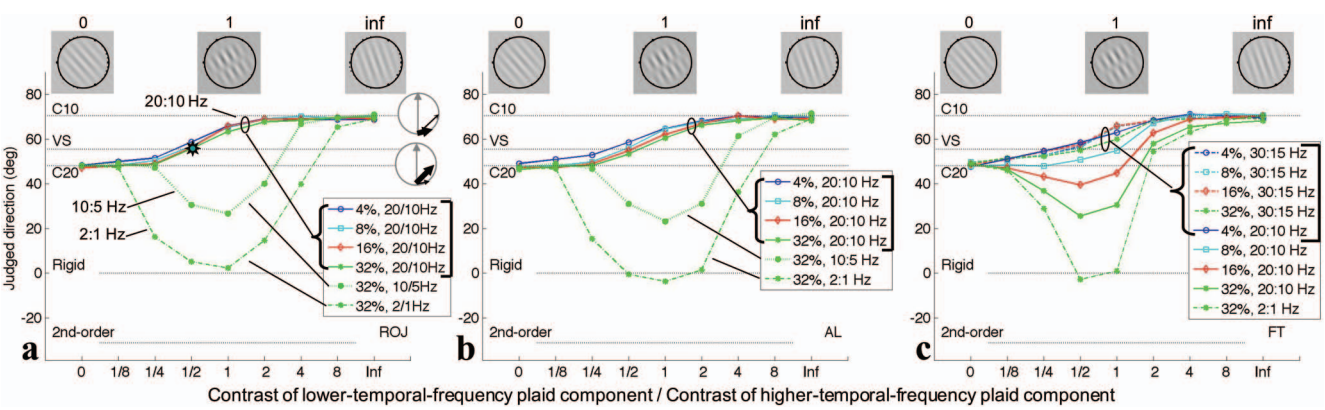


Figure 6. Judged directions of randomly oriented Type II plaids as a function of the contrast ratio and temporal frequencies of the components; data for three subjects. The three thumbnail images above each panel illustrate the two pure sine-wave stimuli and the plaid with equal-contrast components. For each condition, the maximum contrast and temporal frequencies of the components are indicated directly in the panels. The horizontal lines labeled C10 and C20 represent the directions of these plaid components. VS is the direction of the velocity vector sum of the plaid components; Rigid is the direction of rigid translation; 2nd-order is the direction of the strongest second-order motion component; none of these depend on relative contrast, all are indicated as horizontal lines. At temporal frequencies sufficiently high to bypass the third-order system (20 Hz,10 Hz [Panels a, b]; 30 Hz,15 Hz [Panel c]), the judged directions of the four stimulus sets of different contrasts lie on top of each other within measurement error. The judged direction of pure first-order plaid motion depends only on the contrast ratio of the components and is independent of the absolute values of the contrasts over the full range of contrasts. Lower temporal frequencies admit contributions of third-order motion processing (in the direction of rigid translation) to judged motion direction. See text for details and see the online supplemental materials for video examples of the above plaids. See the online article for the color version of this figure.

(but not other contrast ratios) as moving precisely in the rigid direction (pure third-order).

Two other subjects (data not shown) were tested in Phase 1 of the experiment. Their data essentially replicated that of the subjects in Figure 6a and 6b.

The derivation of the motion direction signaled by the second-order system for the various plaid stimuli (horizontal line at bottom of Figure 5) is treated in detail in Experiment 3. The data of Experiment 2 provide no indication that second-order motion plays any role in the perception of the motion direction of these plaid stimuli.

Calculating Contrast-Strength Vectors

Implicit Contrast-Strength Vectors

The fact that equal ratios of component contrasts (independent of the overall contrast level) produce identical perceptions of motion direction means that an underlying power law can describe the data. In the vector sum of velocities algorithm, the perceived motion direction is predicted by the vector sum of the component velocity vectors. In the contrast-strength-vector summation theory, "velocity" is replaced with "contrast strength." Unlike velocity, contrast strength cannot be measured physically, only psychophysically. In thinking about motion, we normally think of velocity, the combination of direction and speed. However, in early vision, many features are computed independently. For example, color, form, motion, and depth seem to be computed in different specialized brain areas, and within these computational areas, different scales (levels of image resolution) are treated separately. Thus, it is not unreasonable to assume that motion speed and direction may be computed separately in the early stages of visual processing (see the "Neural economy" subsection). The computation of exclusive concern here is only perceived motion direction, not speed.

F7 The vector diagram in Figure 7a illustrates how the contrast-strength vector of a plaid component grating is determined. It is assumed that each sine-wave component of a plaid stimulus can be represented as a vector that has a direction—that of the sine-wave motion—and a magnitude—determined by its contrast and temporal frequency. AQ: 27 Figure 7a shows a particular example from Experiment 2 to illustrate the contrast-strength-vector summation theory. It shows a plaid composed of two 1 cpd sine-wave gratings; one with 10-Hz, 32% contrast, the other with 20 Hz, 32% contrast. The angle between the components is $70.5° - 48.2° = 22.3°$. For this particular stimulus, a perceived motion direction was empirically observed. If we consider these two components as motion vectors of equal strength, the addition of these vectors does not produce a resultant that matches the judged direction of the plaid. However, the geometric construction in Figure 7a shows how to produce a shorter vector $\vec{\mathbf{S}}_{20\text{Hz},32\%}$ in the same direction as the component vector in the 20-Hz direction, which, when added to the 10-Hz component vector, does produce the perceived direction. The length $S_{20\text{Hz},,32\%}$ of the vector $\vec{\mathbf{S}}_{20\text{Hz},32\%}$, which is less than 1, is the contrast strength of the 20-Hz 32% grating component.

In general, the two components in any plaid can be compared to determine their relative strength. When there is a strength standard with unit length, the relative strengths become absolute strengths (relative to the standard). So, we begin by assuming that the strongest motion component in these experiments, 10 Hz, 32% contrast, has a contrast strength of 1.0. Using a simple algebraic formulation of the estimation procedure described above, the contrast strengths of all the motion components in both Experiments 1 and 2 were estimated. The same motion component may occur in four different plaids in Experiment 2. A single, average contrast strength was estimated that minimized the sum of squared prediction errors of perceived direction.

The results of the contrast strength estimations are displayed in Figure 7b. This shows the contrast strengths of the sine-wave components of plaids as a function of their contrasts with temporal frequency as a parameter. Only plaid-component sine waves that were processed exclusively by the first-order motion system (10 Hz and 20 Hz for subjects AL and ROJ, and of 15 Hz and 30 Hz for subject FT), and because the logarithms of strength are graphed, only plaid components that are significantly different from zero can be included in the contrast-strength estimates. The straight lines that are fit to the contrast strengths of component sine waves are constrained to have the same slope β_i for both of each subject i's component temporal frequencies. On a log-log plot, the straight lines represent power laws (contrast strength increases as power of contrast) with exponents β_i for subjects FT, AL, ROJ, of 1.6, 2.0, and 2.4. AQ:28, 29

In Figure 7, the antilog of the difference between the 10-Hz and 20-Hz curves for a subject i is the ratio of their effectiveness for motion perception,

$$\Delta = (\text{strength of 20 Hz})/(\text{strength of 10 Hz}) = (0.45, 0.53, 0.33), \quad (1)$$

for the three subjects. The fact that the distance between the 10-Hz and 20-Hz data is independent of contrast means that the single attenuation factor (Δ) applies equal to all the 20-Hz stimuli in the experiment. The power-law fits account for 99% of the variance of the data for the five contrasts from 2% to 32%. The extremely good fits of a power law to first-order motion data confirms the observation (e.g., Figure 6) that the first-order system computes the motion direction of a plaid grating based on the contrast ratio of the component gratings—independent of their absolute contrasts.

Power laws are fairly common in psychophysics. Stevens (1957) AQ: 30 found that the human experience of intensity in 21 sensory dimensions (mostly as judged by magnitude estimation) is well described by a power law with only the exponent changing between dimensions (Poulton, 1967). The amplitude of neural responses to motion stimuli is occasionally found to follow a power law (e.g., Carandini, Heeger, & Movshon, 1997), but it is not necessary that a psychophysical power law requires a power-law neuron amplitude. For example, for more than a century, it was mistakenly thought that Weber law sensitivity required a logarithmic representation of intensity, whereas a feed-forward gain control mechanism provides a much better explanation (Sperling, 1989). How a AQ: 31 functional power law is represented in the brain (and vice versa) has not been determined. A power-law representation of stimulus AQ: 32 intensity near threshold (e.g., a soft threshold) has been invoked to account for the "dipper" effect—given a threshold stimulus x, it is easier to discriminate $2x$ from x than x from 0 (for references, see Appelbaum, Lu, & Sperling, 2007, p. 1). That is, the size of a AQ: 33 threshold increment decreases as the size of the base increment

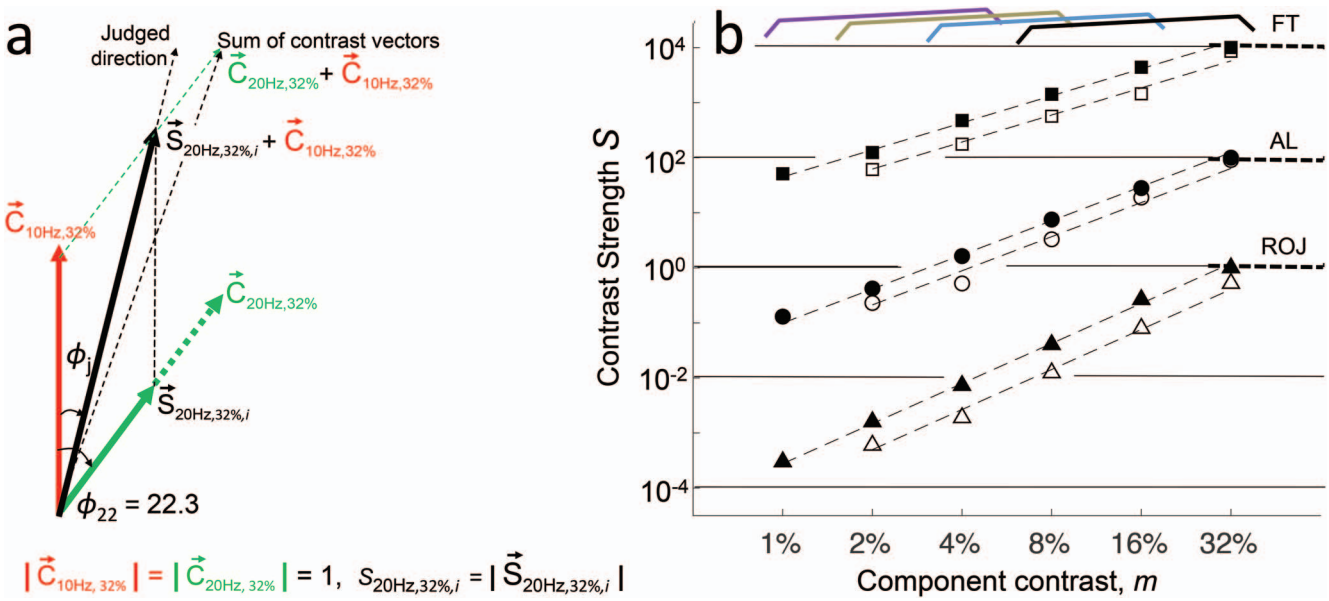


Figure 7. The contrast-strength-vector summation theory and the resulting power law: Contrast strength is a power function of stimulus contrast. (a) Estimating the contrast strength S of a plaid component from the plaid's judged direction. Each plaid component is represented as a vector $\vec{\mathbf{C}}_{h,m}$ that has a direction equal to the direction of the component sine wave and a length proportional to its contrast. In this example, both plaid components have a spatial frequency of 1 cpd and a contrast of 32%. Component 1, $\vec{\mathbf{C}}_{10Hz,32\%}$, has a temporal frequency of 10 Hz and moves directly upward; Component 2, $\vec{\mathbf{C}}_{20Hz,32\%}$, has a temporal frequency of 20 Hz and moves at an angle of $\phi_{22} = 22.3°$ relative to Component 1, as in Experiment 2. The vector sum $\vec{\mathbf{C}}_{10Hz,32\%} + \vec{\mathbf{C}}_{20Hz,32\%}$ of these equal contrast vectors lies on a direction $\phi_{22}/2$ exactly between the two components. However, the judged direction is $\phi_j < \phi_{22}/2$. The geometric construction in Panel a shows that if the length of $\vec{\mathbf{C}}_{20Hz,32\%}$ were reduced to produce a smaller vector $\vec{\mathbf{S}}_{20Hz,32\%}$, in the same direction, the revised vector sum would match the judged direction. The length $S_{20Hz,32\%} = |\vec{\mathbf{S}}_{20Hz,32\%}|$ of the vector $\vec{\mathbf{S}}_{20Hz,32\%}$ is the "strength" of $\vec{\mathbf{C}}_{20Hz,32\%}$. This geometric construction enables estimation of the strength S of Component 2 relative to the strength of Component 1 of any plaid whose perceived direction has been determined. (b) The contrast strength S of the components of the Type II plaid stimuli of Experiment 2 derived from motion-direction judgments by the contrast-strength-vector summation method described above for each temporal frequency, contrast, and subject. The ordinate represents the contrast strengths for subject ROJ. For clarity, contrast strengths are translated upward by two log_{10} units for subject AL and four log_{10} units for subject FT. Filled dots indicate 10-Hz (subjects ROJ, AL) and 15-Hz (FT) gratings. Unfilled dots indicate 20-Hz (subjects ROJ, AL) and 30-Hz (FT) gratings. The straight lines are least squares best fits to the data; they are constrained to have the same slope $\beta_i \approx 2$ for all data from a subject i; they represent power laws with exponents β_i, and they account for 99% of the data variance. The brackets on top indicate the four ranges of stimuli within which useful strength comparisons were possible; they correspond to data within four panels (Panels c, d, e, and f) of Figure 5. See the online article for the color version of this figure.

increases above zero before the increment threshold inevitably increases toward a Weber fraction. The 32:1 range of precise power-law description of behavior in Figure 7 is the result of four overlapping ranges of stimuli. The typical largest useful range within a single stimulus was 8:1.

All the high-temporal-frequency (first-order) stimuli were presented in a mixed list. Thus, the adaptation that enabled the visual system to accurately code the range of contrasts that were contained within a particular stimulus must have occurred during the 200 ms during which the stimulus was presented. How the visual system can instantly adapt automatically and effectively to such different contrast ranges is itself a very provocative question.

Beyond deriving a power law, to be useful, the contrast-strength-vector summation theory must apply in other contexts. Experiment 3 demonstrates this. To understand Experiment 3, we first need to explicitly define what is meant here by different motion systems.

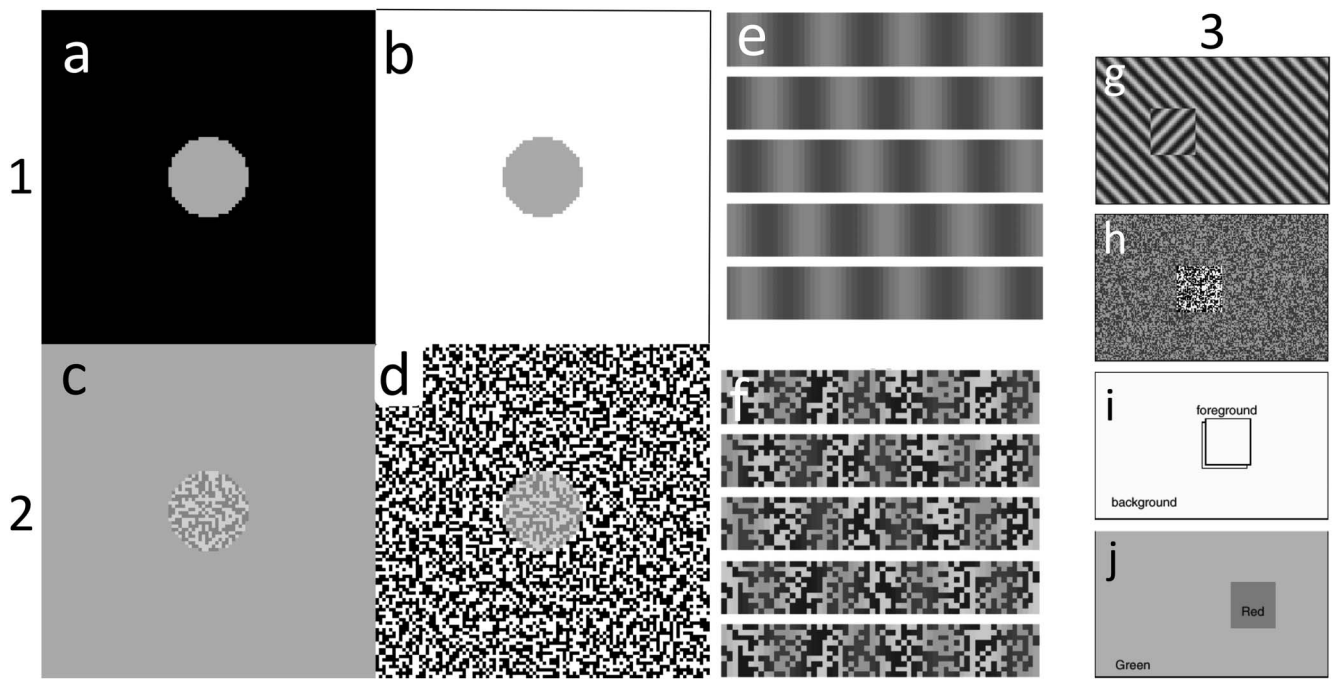
Motion Systems Defined

The motion systems that we are concerned with here are computations that take a scalar function f(x, y, t), for example, luminance as function of space and time, and convert it into a vector flow field g(x, y, t) that is also a function of space and time. The motion vectors of the flow field have a direction and a (nonnegative) magnitude; what the vector magnitude represents will be determined here for the first-order motion system. There are, of course, more complex motion processes that operate on motion

flow fields to compute heading direction, structure from motion, balance, and so on; these are not under consideration here. The aspect of the first-, second-, and third-order motion computations that is of concern here is the computation of motion direction. The differences between these motion computations are best understood in terms of the preprocessing of the input prior to the motion computation rather than in terms of the differences in the motion computation itself.

F8 Figure 8 illustrates stimuli that are exclusively directed to each one of the three orders of perceptual processing. Common nomenclature is that luminance-modulation stimuli are directed to the first-order motion system and that contrast-modulation stimuli are directed to a second-order or, sometimes, simply a higher order AQ:35 motion system. It is misleading to think of luminance as the input to a motion computation. Luminance is a non-negative quantity, whereas an essential aspect of first-order motion is that sign matters, that an input to first-order motion is either positive (a point has more luminance than its surround) or it is negative (darker than its surround). Reichardt (1961) demonstrated the phenomenon he called "correlation," subsequently called "reverse phi" by Anstis (1970): Two successive adjacent positive flashes appear to move in the same direction as two successive adjacent negative flashes; but when successive flashes have different signs, positive and then negative, or vice versa, they can appear move in a direction opposite to that of same-sign flash pairs.

The visual preprocessing prior to the first-order motion computation is much more complicated than merely computing Weber contrast (i.e., luminance adaptation): It involves spatial filtering



AQ: 70 *Figure 8.* Examples of first- and second-order spatial interactions and of prototypical stimuli directed to the three perceptual systems for discriminating motion direction. (a, b) First-order: The center disk has equal physical contrast in Panels a and b. It appears darker in Panel b than Panel a, because in Panel b, it is surrounded by a higher mean luminance. (c, d) Second-order: The center disk in Panels c and d has the same physical texture contrast. It appears to be of lower contrast in Panel d than Panel c, because in Panel d, it is surrounded by a greater texture contrast (higher variance texture). (e) First-order motion: The sine wave in each of the five frames in this panel differs from the adjacent frame by 90°. When the five frames are presented successively (from top to bottom), superimposed, Panel e represents a "luminance" motion stimulus moving to the right directed to the first-order motion system. Alternatively, this panel, as illustrated, represents a first-order sine-wave grating slanted to the right. (f) Second-order motion: Each of the frames in this panel represents a sine-wave modulation of texture contrast. The contrast modulation in each of the five frames in this panel differs from the adjacent frame by 90°. Presented successively (from top to bottom), superimposed, this panel represents a "contrast" motion stimulus moving to the right directed to the second-order motion system. Alternatively, this panel, as illustrated, represents a second-order sine-wave grating slanted to the right. (g–j) Third-order motion: Four frames in which the central rectangle ("figure") is differentiated from its surround ("ground") by, successively, (g) slant, (h) contrast, (i) stereo-depth, and (j) color. Presented successively, superimposed, the display is perceived as a square that changes its substance from frame to frame as it moves from left to right. This motion example is merely illustrative; actual experimental stimuli are more complex. Panels [a–d after Figure 1 Chubb, Sperling, and Solomon (1989), p. 9632; Copyright (1989) George Sperling, by permission; panels e, f after Figure 7, panels g–j after Figure 4, Lu & Sperling, 2001, p. 2338 and p. 2335, Optical Society of America, by permission.]

into spatial frequency channels, orientation filtering within channels (Blakemore & Campbell, 1969), light adaptation, and contrast-gain control (Lu & Sperling, 1996), all before the motion computation.

The computations in second-order motion (as defined by Lu & Sperling, 1995a, 2001) involve all the processes listed for first-order motion, except that second-order contrast-gain control is different from first-order (Lu & Sperling, 1996). However, the main difference is that in second-order motion, sign does not matter—positive and negative deviations from the mean level are treated approximately equally. Opposite-sign successive adjacent flashes and same-sign adjacent flashes produce the same direction of movement in second-order motion processing. Sign indifference is represented as a rectifying operation, that is, taking the absolute value or the square of the input variable. Using a statistical analogy, first-order motion operates on the local mean value of the input, and second-order motion operates on the local variance of the input. Of the three motion computations considered here, second order is by far the weakest.

The preprocessing prior to third-order motion (as defined by Lu AQ: 72 & Sperling, 1995b, 2001) is best understood as the computation of the motion of figures in a figure-ground representation. More formally, the output of the figure-ground process is represented in a salience field, in which "figure" areas have higher values of salience than "ground" areas. Whereas figure-ground is often thought of as a binary variable, salience is a continuous variable.

Third-order motion is often referred to as *feature tracking*; however, this is a misconception. To track something first requires a motion computation to give a direction for tracking. Various studies (Blaser et al., 1999; Lu & Sperling, 1995b; Tseng et al., 2004) demonstrate that attending to a feature produces higher salience in the regions occupied by the feature, and the third-order motion computes the movement of the salient areas. Once the direction of motion has been computed, a feature can then be tracked by attention with or without eye movements. Without a motion computation, tracking a feature would require a search process to discover where it had moved. Although much is known about the spatial and temporal resolution of third-order motion for certain classes of stimuli (Lu & Sperling, 2001), the full details of the preprocessing for salience AQ: 36 have not yet been worked out. What is clear, however, is that third-order motion is inherently concerned with the motion of perceived objects, that is, area of salience. The first-order system does not know about objects—it merely computes local flow fields. Relating first-order motion flow fields to perceptual objects requires a subsequent "binding" process.

Second-Order Motion Contributions to Plaid Motion

A problem in defining second-order motion is that the word *contrast* is used for two different concepts: global and local. *Global* (or *texture*) *contrast* refers to the square root of the variance of a texture or a sine wave. *Local* (or *point* or *Weber*) *contrast* refers to the difference of point from its surround. Global contrast is non-negative; local contrast is as likely to be positive as negative. The second-order motion system ignores the sign of the local contrast, that is, the input to second-order motion processing is a function of the absolute value of the local contrast. This function can simply be the absolute value or the square, or some more complex function. Here, we assume that the input to second-order motion it is the square of the local contrast. All of the square computations were repeated with absolute value. For the plaids under consideration, there were no significant differences in computed motion direction between the squared- and absolute-value plaids.[1] Fn1

Figure 9a illustrates a snapshot of a plaid stimulus, Figure 9b F9 illustrates the squared contrast of the stimulus, and Figure 9c illustrates a vector diagram illustrating the direction of motion of the dominant sine-wave component of the squared stimulus as well as the directions of the original sine-wave components in which the direction of motion of a pure sine wave is taken as perpendicular to the orientation of the sine wave. In this particular example of a Type II plaid stimulus from Experiment 2, the directions of second-order motion, the direction of rigid motion, and the direction of motion perpendicular to the perceived orientation of the plaid pattern are all widely separated. In the case of squaring the point contrasts to produce the second-order motion input, the directions of the resulting motion components can easily be calculated. However, the amplitudes of squared components cannot be specified relative to the first-order sine wave because they depend on the units of measurement; squared amplitudes can be specified for a display but not for the stimulus representation in the nervous system. On the other hand, when the absolute value—versus the square—of point contrast is used to produce the second-order motion input, the amplitude of second-order component can be specified relative to that of the first-order components. In the AQ: 38 case of an equal-contrast, 10 Hz and 20 Hz plaid with a grating component frequency of one cycle per degree and an angle of 22.3° between the components, the amplitude of the strongest second-order component is 0.50 times the amplitude of the original stimulus components. The second-order component in the stimuli of Experiment 2 has a large stimulus amplitude. In Figure 6, the horizontal line near the bottom of the panels illustrates the direction of the dominant second-order component. There is no indication that, for these plaids, the second-order component has any influence on judged motion direction.

Experiment 3: Perceived Motion Direction as a Function of the Angle Between Plaid Components

For any single-spatial-frequency plaid that stimulates only the first-order motion system, the contrast-strength-vector summation theory predicts the perceived direction of plaid motion. To fully exploit this theory, consider two plaid components (sine waves) that have the same spatial frequency, and possibly different temporal frequencies, f_1, f_2, contrasts, c_1, c_2, and strengths, S1(f_1, c_1), S2(f_2, c_2). Assume that the relative strength $\Delta = S2/S1$ is known, for example, because it was determined in a prior experiment. Assuming the temporal frequencies are sufficiently high and/or the contrasts are sufficiently low to guarantee that S1 and S2 stimulate only the first-order motion system, Δ is sufficient to enable pre-

[1] The direction of rigid translation of a plaid stimulus is the same whether computed on the original plaid pattern or on any transformation of the plaid pattern that leaves it two-dimensional. Because any plaid stimulus can be represented as the translation of a snapshot of a plaid, the direction of rigid translation does not depend on how the plaid is represented in the snapshot (Figure 2).

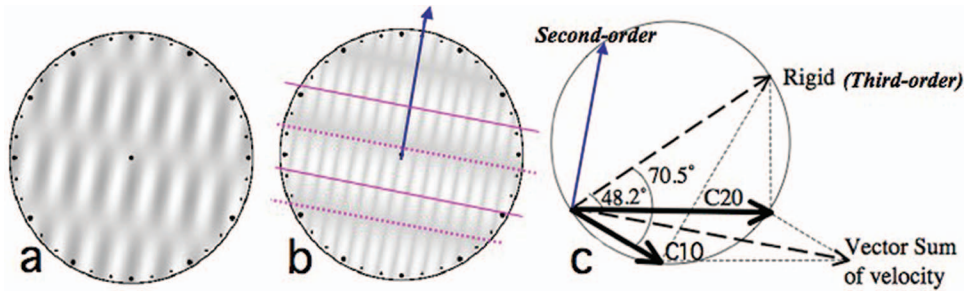


Figure 9. Motion direction of a Type II plaid according to a second-order motion computation. (a) Snapshot of a Type II plaid. (b) The same stimulus as Panel a except that the point-contrast values have been squared (rectified). The two solid lines in Panel b indicate the positive peaks of the fundamental sine wave—the dominant sine-wave component—in the rectified stimulus pattern. The two dotted lines indicate the negative peaks of the dominant sine-wave component. The arrow indicates the motion direction of the dominant second-order sine wave. (c) A vector diagram showing the motion directions of the component sine gratings (C10, C20), the rigid direction, the velocity vector sum, and the direction of second-order motion. See the online article for the color version of this figure.

diction of the perceived direction of any plaid composed of components S1 and S2 no matter what angle between them may be. For any new plaid, the direction of the sum of the components' contrast-strength vectors predicts the perceived motion direction, and Δ is sufficient to determine their relative strengths.

Method

AQ: 39 To test the contrast-strength-vector summation theory, Experiment 3 takes two different pairs of gratings from Experiment 2 whose relative contrast strength has been determined in a Type II plaid with an angle of 22.3° between the components. For each of these pairs, Experiment 3 tests the perceived directions of plaids composed of these components over a wide range of angles between components. If the contrast-strength-vector summation theory is correct, the outcome of Experiment 3 will be perfectly predicted from the 22.3° observations in Experiment 2, with zero parameters to be estimated in Experiment 3.

In Experiment 2, the angle between the two components of the plaid was 22.3°, and contrast and temporal frequency were varied. Here, six angles between 20° and 120° were used (0° and 180° are not useful), and the spatiotemporal frequencies of the components were the same as those used in Experiment 2: 1 cpd, 10 Hz and 20 Hz. Two pairs of contrasts were used: In both pairs, the 20-Hz grating had a contrast of 4%; in one plaid, the contrast of the 10-Hz grating was also 4%, and in the other, it was 2%. As in Experiment 2, the rigid direction was randomly chosen on each trial from 0° to 359°, two mirror-image flips of each angle were tested, and the results are combined. Two subjects who participated in Experiment 2 also participated in Experiment 3; one new subject was recruited.

Results

F10 Figure 10 shows the judged motion direction for each of the stimulus plaids for three subjects as a function of the angle between the plaid components. The reference point in Figure 10 is the direction of the 20-Hz component, the lower strength, faster-moving component. As the direction of the 10-Hz component deviates more and more from the 20-Hz component, the perceived direction of the plaid moves further and further from the 20-Hz direction toward the 10-Hz direction but always remains in between. The movement of the data away from the 20-Hz direction is obviously greater for the stronger 4% contrast 10-Hz component than for the 2% contrast 10-Hz component. For the two subjects AQ: 40 who participated in Experiment 2, the lines through the data points use the contrast-strength ratio Δ derived from the 22.3° plaids in that experiment. For the new subject, LL, the constant Δ is derived from the 22.3° plaids in Experiment 3 itself. The same Δ that describes the relative contrast strength of two components of a plaid in Experiment 2 also predicts quite accurately how these components will combine in the seven new plaids of Experiment 3.

In Figure 10, for the two subjects who participated in Experi- AQ: 41 ment 2, the predicted curves for both types of plaids (component contrasts of 4%, 4% and 2%, 2%) account for over 97% of the variance of the data. Zero parameters were estimated from Experiment 3 itself; that is pure prediction. In contrast, the vector sum of velocity, the direction of rigid translation, and the motion direction of the primary second-order motion component make identical predictions for both sets of data, and the predictions are far off the mark and, obviously, do not account for the data.

Conclusion

For a plaid grating that moves with temporal frequencies so high and contrasts so low that only the first-order motion system contributes to perception, only one parameter, the relative contrast-strength Δ of the two sine-wave components, is sufficient to enable almost perfect prediction of judged motion direction.

Experiment 4: Orientation Judgments Versus Motion-Direction Judgments—A Control Experiment

All plaid patterns used in this study have predominant orientations. Therefore, in principle, subjects could use the predominant orientation of a static plaid as a clue to infer the perceived direction of the moving plaid. Orientation is defined by a snapshot of the

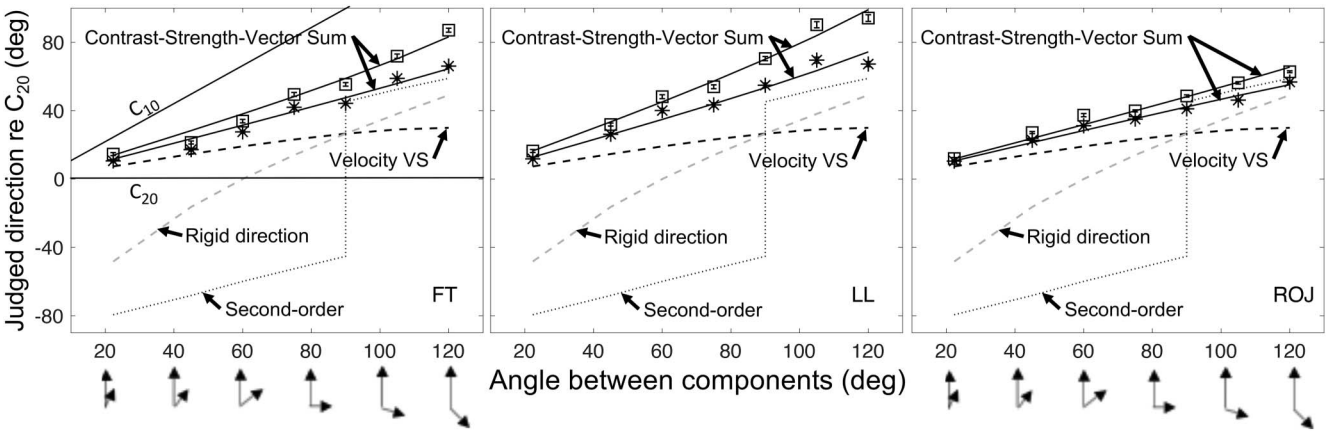


Figure 10. Predicted and perceived directions of randomly oriented plaids as the angle between the component gratings is varied. Each panel represents data points from one subject plus five continuous curves representing theoretical predictions of four theories. As in Experiment 2, one grating component of the plaid has a temporal frequency of 20 Hz, and the other, 10 Hz; the contrasts of the data points are as follows: squares (C20, 4%; C10, 4%); asterisks (C20, 4%; C10, 2%). The abscissa indicates the angle between the components and illustrates the configuration; the longer vertical arrow is the 20-Hz component. The ordinate indicates the judged motion direction relative to the direction of 20-Hz grating. The two uppermost curves are the a priori predictions of motion direction based on the contrast-strength-vector sum of the two plaid component gratings using the relative contrast-strength constants Δ (Equation 1) from Experiment 2 (for subjects FT and ROJ) and a Δ estimated from only the 20° data for a new subject LL. The a priori predictions account for over 97% of the variance of the data for both conditions for each subject. Also shown, the predictions for vector sum of velocities, rigid direction, and the direction of second-order motion make identical predictions for both sets of data. The direction of second-order motion flips 90° as the angle between the gratings changes from 89° to 91°, and the strength of second-order motion approaches zero as the angle between the gratings approaches 0° or 180°. The solid straight lines (left panel only) show the representation in this graph of the directions of the plaid components C20 and C10. See the online article for the color version of this figure.

F11 plaid (see Figure 11). Orientation judgments cannot be used exclusively to judge motion direction because a snapshot of a plaid is the same, independent of the temporal frequencies of the components that determine a wide range of different rigid (objective) and perceived motion directions. For this Type II plaid, the perceived motion direction is between the C10 and C2 components at high temporal frequencies and in the direction of rigid translation at very low temporal frequencies. The influence of orientation per se is obviously limited. Nevertheless, to evaluate the potential confounding factor of orientation versus the factor interest, motion, for stimuli like those used in the main experiment, orientation judgments and motion-direction judgments were measured and compared. The principles are illustrated in Figure 11. The method and results are fully described in the online supplemental materi-
AQ: 42 als. The main conclusion is that orientation judgments have a much smaller variance than motion direction judgments, and a much smaller variation with intermediate contrast ratios of the components, than do motion judgments.

Plaid Motion: Theory

Summary of Results To Be Explained

The experimental results that a theory of plaid motion perception must account for are summarized here. A critical preliminary result was that the motion stimuli themselves must not be concentrated around a particular direction—that led to biased direction estimates. The present procedure, in which the physical motion direction of the stimulus was randomly selected from a 0° to 359° range of directions, led to unbiased judgments of perceived motion direction.

The main experiment, Experiment 2, investigated Type II plaids composed of two 1-cpd gratings. One component moved in a direction 70.5° from the rigid direction; the second component with twice the temporal frequency of the first moved 48.2° relative to the rigid direction. For temporal frequencies of 10 Hz and 20 Hz (two subjects) and 15 Hz and 30 Hz (one subject), and for contrasts of the higher contrast component ranging from 2% to 32%, the perceived plaid direction was entirely determined by the contrast ratio of the components independent of their overall contrast. For example, a grating with component contrasts of 4% and 2% appears to move in precisely the same direction as a similar grating with contrasts of 32% and 16%. For equal contrasts of the two components, the perceived direction of the four Type II plaids (contrast = 32%, 16%, 8%, 4%) was 63° ± 2° (relative to the rigid direction) for the three subjects. When the contrast of stronger component exceeded that of the other by a factor of 4 or 8, the lower contrast component was usually ineffectual, and perceived direction was the direction of the higher contrast component (either 48.2° or 70.5°). For smaller contrast ratios, intermediate directions between the two components' directions were systematically observed.

When component temporal frequencies were decreased by 5 Hz and 10 Hz (two subjects), and 10 Hz and 20 Hz (one subject), into

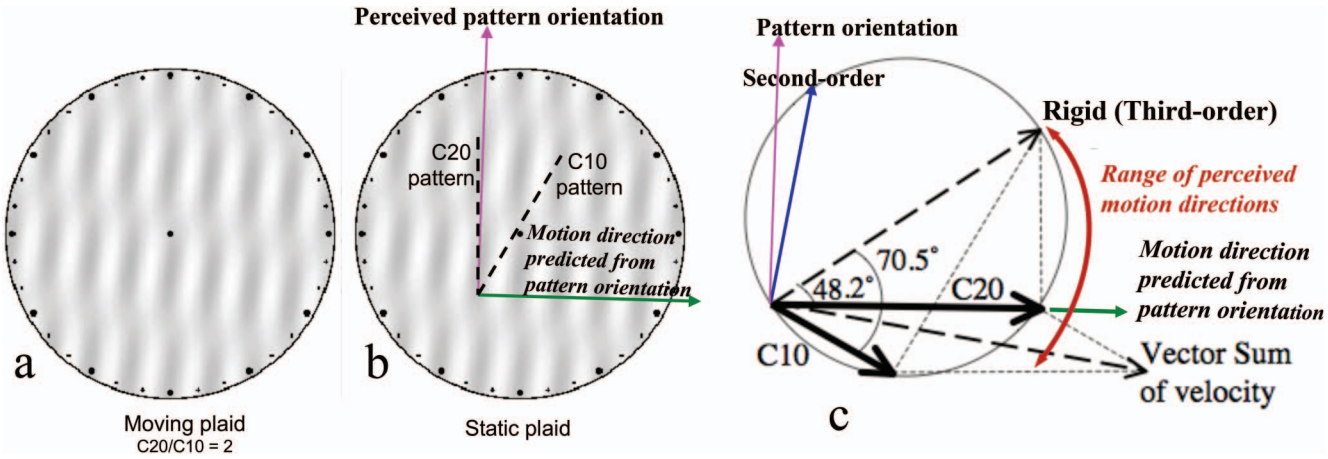


Figure 11. Pattern-orientation judgments in a snapshot of a Type II plaid. (a) Snapshot of a Type II plaid. (b) A pattern-orientation judgment made by a subject. Pattern-orientation judgments seem to be based on the orientations of the component micropatterns. To compare a pattern-orientation judgment with a motion-direction judgment, the motion direction is assumed to be perpendicular to the pattern orientation. (c) The various judgments and theories compared. For this Type II plaid, the perceived motion direction depends on temporal frequency. For high temporal frequencies (the range investigated in Experiment 4), perceived motion direction lies between C10 and C20. See the online article for the color version of this figure.

temporal frequency ranges in which third-order motion was possible, the perceived direction of plaids with equal or near-equal component contrasts, particularly high absolute contrasts, deviated increasingly toward the rigid direction as temporal frequency was reduced. For our Type II plaids with component temporal frequencies of 1 Hz and 2 Hz and contrasts of 32%, all subjects perceived the plaid as moving precisely in the rigid direction, that is, pure third-order motion with no measurable influence of first-order motion.

Using a simple contrast-strength vector model, the relative contrast strength Δ of the two component gratings of a plaid was derived from the perceived direction of the plaid. For the three subjects, the contrast strength of a sine-wave plaid component was found to be a power of its contrast (power exponent $= 2.0 \pm 0.4$ for the three subjects). A power law is consistent with direction being determined by the ratio of component contrasts, independent of the overall contrast magnitudes. To test whether this derived strength measure describes the behavior of the same components in new combinations, the angle between the two component gratings of a plaid was systematically varied. A single parameter derived from Experiment 2, relative contrast strength Δ, accurately described the perceived direction of two-component plaids for all angles between the components tested in Experiment 3. In the next sections, we consider how various prior theories of plaid motion can (or cannot) deal with these basic findings, and we propose a new theory.

Prior Theories of Plaid Motion

The plaid stimulus was introduced by Adelson and Movshon (1982). They proposed a geometrical construction for finding the direction of rigid translation using an IOC algorithm (e.g., Figure 1). However, for plaids with unequal contrasts, the perceived direction was biased toward the higher contrast component (Stone et al., 1990). However, the rigid direction is independent of contrast. As a result, Stone et al. (1990) proposed an IOC of perceived speed model. Nevertheless, neither IOC of physical component speeds nor of perceived component speeds could account for perceived plaid motion direction when the perceived speeds were experimentally measured instead of by model fitting (Champion et al., 2007). Therefore, Champion et al. (2007) proposed that Bowns (1996) feature tracking explanation that motion of the zero-crossing edges of the features might contribute to the perceived direction. Feature tracking (i.e., third-order motion) indeed occurs, but only at temporal frequencies within the range of a subject's third-order motion system. In the present experiments, at sufficiently high temporal frequencies to exclude third-order motion (i.e., above 10 Hz, two subjects; 15 Hz, one subject), feature tracking was excluded and the perceived motion direction was determined entirely by the relative contrast strength of the plaid components. AQ: 43 AQ: 44

Fleet, Black, Yacoob, and Jepson (2000) proposed linear parameterized affine models of optical flow for computing velocity (speed and direction) of complex motion stimuli. The effect of contrast was not considered in these models. As contrast is critical for both first-order and third-order motion processes, this model is not appropriate for human vision.

A vector summation model that combines Fourier and non-Fourier (second-order) component motion signals was proposed by Wilson et al. (1992), Wilson and Kim (1994), and Yo and Wilson (1992). The most serious problem with this model is that, in our plaid stimuli, there was no evidence of any contribution of second-order motion to perceived direction when using the same second-order computation as the Wilson group.

Weiss et al. (2002) developed an optimal Bayesian estimator that considers effect of contrast. In their model, for a low-contrast grating, the likelihood is more spread out (more noise/uncertainty);

the prior is that velocities tend to be slow. They used IOC (direction of rigid motion) to describe the visually computed motion direction prior to noise uncertainty. Their model obviously deals only with third-order motion because, as was demonstrated with the Type II plaids in Experiment 2, for temporal frequencies above 10 Hz to 15 Hz, the direction of perceived motion always lies between the two components' directions and never in the rigid direction, which was 48° away from the nearest component direction.

Another group of theories that are potentially relevant to plaid perception deal with how different local motion signals are combined across spatial areas (Amano, Edwards, Badcock, & Nishida, 2009; Maruya, Amano, & Nishida, 2010; Mather, Pavan, Bellacosa Marotti, Campana, & Casco, 2013; Sun, Chubb, & Sperling, 2014, 2015). In particular, Amano et al. (2009) showed subjects arrays of nonoverlapping translating Gabor patterns. When each half of all of the Gabor patterns translated in a direction with a temporal frequency that was consistent with a Type II configuration, two of three subjects showed a bias toward the rigid direction. This result is interesting because nonoverlapping Gabor patterns do not produce features that move in the rigid direction, so any perception in the rigid direct direction must have been produced otherwise, possibly by an IOC computation. However, unlike Experiment 2 (above), in which all subjects observed 100% perfect rigid direction perception, Amano et al.'s two of three subjects observed only a weak bias toward the rigid direction. This requires further investigation before reaching a firm conclusion about mechanism.

So, we are left, like Sun et al. (2015), with a theory in which first- and third-order motion combine their outputs to produce a final, common, perceived-motion-direction output. This theory, as applied to plaids, is formalized below.

Considerations for a Theory of Plaid Motion

Neural economy: An a priori principle for independent direction and speed computations. A very simple principle of early vision is the breakdown of the visual stimulus into separate spatial frequency channels, and within channels, separate computation of motion, depth, color, texture, shape, and other stimulus properties. A simple rational for separate systems for the computation and motion direction and of speed is neural economy. Suppose we wish to represent 10 directions and 10 speeds of motion. If a single neuron was to represent both a direction and a speed, that is, a velocity neuron, then $10 \times 10 = 100$ neurons would be required to represent the 100 possible combinations. If the direction and speed were represented separately, only 10 speed and 10 direction neurons are required, $10 + 10 = 20$, a very substantial saving. When there are n dimensions being represented, the savings of $10n$ versus 10^n is enormous. The separate representation of direction and speed at early stages of processing does not resolve the problem that velocity may have to be computed at some later stage. However, the visual system seems to have been designed to postpone feature combination as long as possible in order to maintain neural economy for a long as possible.

Motion models output the square of input contrast amplitude. There are three more or less equivalent models of early visual motion-direction processing (Adelson & Bergen, 1985; van Santen & Sperling, 1984; Watson & Ahumada, 1985) that all are equivalent for reasonable ranges of their parameters (van Santen & Sperling, 1985), and therefore all account for the six counterintuitive results reported by van Santen and Sperling (1984). The motion detector in each of these models (the motion energy, Reichardt, and Hilbert models) produces spatially localized vectors that represent a direction and the strength of motion in that direction—in which strength is proportional to the square of contrast amplitude. Such detectors only report the contrast strength of motion in a particular direction. They are not directly for speed detection. However, several more elaborate computational schemes have been proposed for deriving speed from arrays of such detectors, for example, Heeger (1987) and Perrone (2012), and from such arrays enhanced with other types of detectors, for example, Burge and Geisler (2015).

Spots, features, boundaries—broadband localized objects—are best for speed detection versus narrowband plaids. Vision psychophysicists and visual neuroscientists have largely studied narrow-band windowed sine-wave stimuli (Gabors) and large clouds of dots drifting in a particular direction with various signal-to-noise ratios. On the other hand, in a classic article, Lettvin, Maturana, McCulloch, and Pitts (1959) reported that what the frog's eye told the frog's brain was the location of dark spots in the visual field. Humans do not appear to have such low-level dot detectors. However, Mach (1886/1959, p. 234) showed that a single dot or several dots placed on an ambiguously moving stimulus immediately determined its perceived direction and shape. What Mach's observations suggest is that, for studying motion direction and even, more particularly, speed, narrow-band stimuli such as plaids do not represent the broadband spots and other local features in the natural environment for which the computation of direction and speed is most efficient. As noted above, at low temporal frequencies, a nonlinear transformation prior to the third-order motion computation can transform the darkest and lightest areas in plaids into salient features, that is, dark areas and light areas. According to the salience theory of third-order motion, it is the movement of the salience produced by these features that enables the computation of the direction of rigid translation. Just as spatially broadband stimuli that maintain a coherent phase relation contain the best information for speed, neurons that are selectively sensitive to speed versus merely to temporal frequency combine outputs over different spatial frequencies (Perrone & Thiele, 2001).

Conclusion. The computations of motion direction and of speed are separate neural and perceptual processes that share common elements; the computation of speed seems to be more complex than the computation of direction. Velocity, that is, speed and direction, is most unambiguously estimated from the movement of a point or of a feature, that is, a broadband input that contains a wide range of sine waves in coherent phase. Insofar as the third-order motion system computes the direction of rigid translation of Type II plaids, it most likely does it by a nonlinear transformation of the plaid into features such as dark and light spots at the locations of the negative and positive contrast maxima.

The Contrast-Strength-Vector Summation Theory

Four assumptions. The contrast-strength-vector summation theory (Figure 12) describes a completely hypothetical, unconscious computation to account for the present data. The theory considers only equal-spatial-frequency plaid stimuli. It takes as

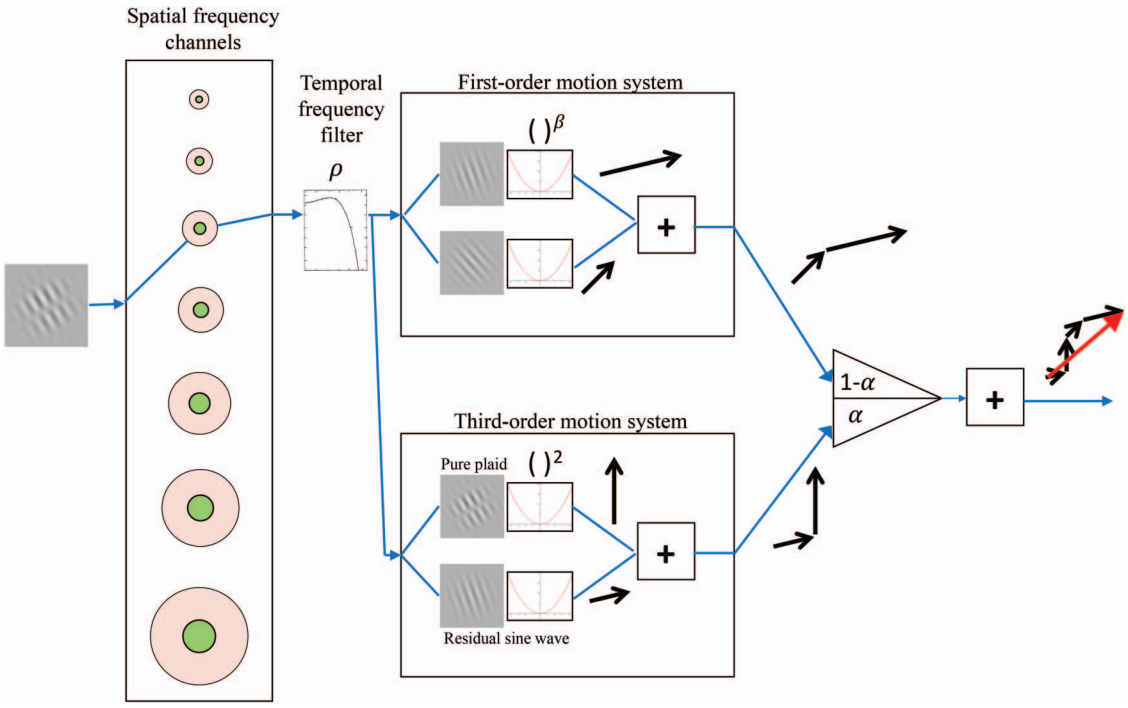


Figure 12. Flowchart of the contrast-strength-vector summation theory to account for the perceived motion direction of same-spatial frequency plaids. The input is a plaid that is processed within a single-spatial-frequency channel. A temporal, lowpass filter attenuates the input contrast of plaid components by a factor ρ_{fi} that depends on frequency *f*, and varies somewhat between subjects, *i*. The temporally filtered plaid is input into both a first-order and a third-order motion system. In the first-order system, each plaid component produces a vector perpendicular to the grating with vector length proportional to a power $\beta \approx 2$ of component contrast, and the two component vectors sum. In the third-order system, one vector is pointed in the direction of rigid translation with length proportional to the squared amplitude of the "perfect plaid" (in which both components have the amplitude of the weaker plaid component). A second vector has the direction of the larger sine-wave component and a length proportional to the square of the residual (the difference between the sine-wave amplitudes). In the present experiments, the relative contribution α of the first- and third-order systems depends on the temporal frequencies and durations of the plaid components. In this illustration, $\alpha = .5$. The output direction is the α-weighted sum of the four vectors computed by the two motion systems. See the online article for the color version of this figure.

input the two plaid sine-wave components that can vary independently in contrast (c_1, c_2), speed (i.e., temporal frequency, f_1, f_2), and angle of motion direction (θ_1, θ_2). This is a large, six-dimensional input space. The theory output is a predicted perceived direction, that is, an angle relative to the rigid direction. The theory deals with two-dimensional motion-direction vectors of the form **V**(length, angle). The four assumptions are as follows:

1. The first-order system processes input plaid components of different angles independently. For each input component, it produces a two-dimensional vector output $\mathbf{C}_j((\rho_j c_j)^\beta, \theta_j)$; ρ_j is the component's temporal-frequency-dependent contrast attenuation ($\rho \leq 1$), c_j is its contrast, θ_j is its motion-direction angle, β is 2 ± 0.4, and $j = 1, 2$. In this and subsequent assumptions subscripts "1, 2" designate plaid components.

AQ: 53

AQ: 54

2. The third-order system's two components are (a) a rigid-direction plaid-component $\mathbf{C}_r((2c_n)^2, 0)$ with amplitude $(2c_n)^2$, where $c_n = \min(\rho_1 c_1, \rho_2 c_2)$ and direction 0; and (b) a sine-wave component $\mathbf{C}_s(||\rho_1 c_1 - \rho_2 c_2||^2, \theta_m)$ with amplitude $||\rho_1 c_1 - \rho_2 c_2||^2$ and direction θ_m, $c_m = \max(c_1, c_2)$.

3. The first- and third-order motion system's outputs are motion-direction vectors **V1** and **V3**, each of which is the vector sum of its system's two component-direction vectors:

$$\mathbf{V1}(L_{V1}, \theta_{V1}) = \text{VectorSum}(\mathbf{C}_1, \mathbf{C}_2), \quad \text{(a)}$$

and

$$\mathbf{V3}(L_{V3}, \theta_{V3}) = \text{VectorSum}(\mathbf{C}_r, \mathbf{C}_s), \quad \text{(b)}$$

where *L* and θ are length and angle of the output vector; subscripts 1 and 2 refer to the plaid components; and subscripts *r* and *s* refer to the perfect plaid (rigid direction) and to the residual sine component.

4. The net predicted direction Θ depends on the frequency-dependent relative strength α of the first- and third-order system outputs, **V1** and **V3**. As the experiments measure

only the direction of judged motion, not its magnitude, the theory assumes that the predicted perceived angle Θ is simply the α-weighted mean of the outputs of the first- and third-order systems:

$$\Theta = (1 - \alpha)\theta_{V1} + \alpha\theta_{V3}.$$

Theory predictions with default parameters. The experiments provide abundant data that enable a very precise theory of the perceived direction of same-temporal-frequency plaids that exclusively stimulate the first-order system, for example, the remarkably linear power-law predictions (see Figure 7) and the predictions for plaids with different component angles (see Figure 10). There are only sparse plaid data that involve the much more complex third-order system. There is no independent confirmation or elaboration of the ad hoc assumption that the third-order motion system regards a plaid as split into a perfect plaid plus residual sine wave or of the assumed square-law combination of these two components. The aim here is to demonstrate that these assumptions are nevertheless sufficient to enable, a priori, a prediction of our entire data set with default parameters.

Three parameters are necessary to predict the data: β, the power-law exponent that determines the relative strength of different plaid component contrasts; ρ, the contrast attenuation of different temporal frequency components; and α, which determines the proportions of first- and third-order motion. A basic property of the Reichardt and similar models is that the strength of sine-wave motion components is proportional to the square of their amplitude. This was empirically confirmed by van Santen and Sperling (1984). Thus, the default β is 2.0. AQ: 55

In a moving plaid the ratio Δ, the strength of the higher temporal frequency component divided by the strength of the lower temporal frequency component, is easily estimated. However, we could not find usable suprathreshold data to estimate motion strengths. We estimate default ρ from the ratio of detection thresholds—a reciprocal measure of strength that, unfortunately, is more influenced by internal noise than the suprathreshold stimuli of the experiments. The motion-threshold-contrast ratio Δ_{th} for 10Hz and 20 sine gratings is (threshold 10Hz)/(threshold 20 Hz). According to Kelly (1979, p. 1342, Figure 3), $\Delta_{th} = 0.244$, which is smaller than the average motion-direction Δ of 0.44 observed in Experiment 2. AQ: 57

The constant alpha, $0 \leq \alpha \leq 1$, determines the relative proportions of first- and third-order motion in the model output. Alpha depends on the plaid component temporal frequencies, on stimulus duration, and on contrast. Figure 13 displays the full range of α from 0 to 1. F13

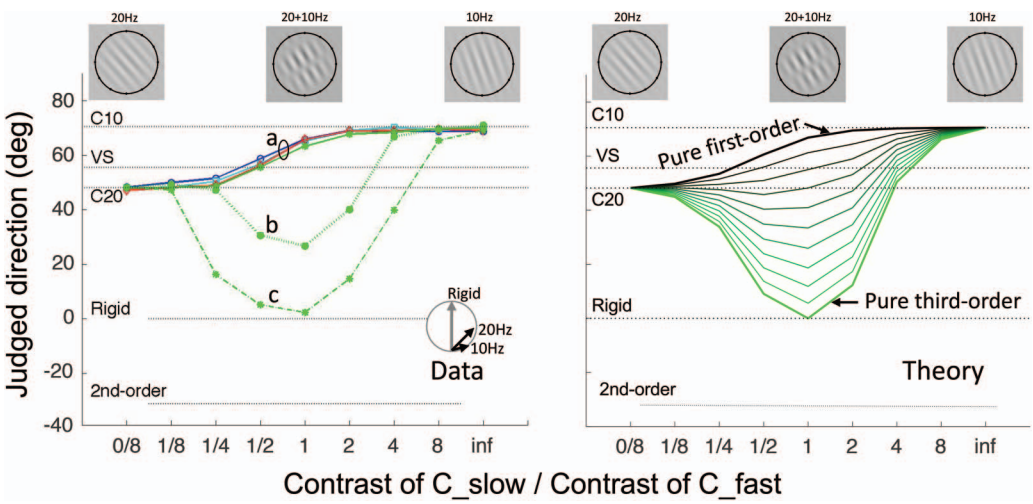


Figure 13. Data and theory for one subject judging the motion direction of Type II plaid stimuli as a function of the contrast ratio of the component sine waves (abscissa) with overall contrast and temporal frequency as the parameters. All component sine gratings were 1 cpd. Data are reproduced from Figure 6a. Ordinates: Data = judged plaid motion direction relative to the direction of rigid translation; Theory = predictions of judged motion direction by the contrast-strength-vector summation theory with default parameters. The horizontal lines labeled C10 and C20 represent the angles (relative to rigid direction) of the 10- and 20-Hz sine-wave components. VS represents vector sum of their velocities. Second-order represents the direction of second-order motion. The direction of rigid translation was varied randomly from 0° to 359°; all angles are given relative to the rigid direction. The insert in the data panel indicates the plaid's rigid direction and the velocity vectors of the two component sine waves. The plaid thumbnails at the top indicate (from left to right): the pure 20-Hz stimulus, a plaid with equal-contrast 10- and 20-Hz components, and the pure 10-Hz stimulus. (a) The curve parameter indicates component frequencies 10 Hz + 20 Hz (200 ms), and four overlapping curves with the contrast of the higher contrast component = 32%, 16%, 8%, and 4%. The perfect overlap of the four curves means that only the contrast ratio matters. (b, c) For these data, the contrast of the higher contrast plaid component was 32%. Component frequencies were, for Panel b, 10 Hz + 5 Hz (650 ms), and for Panel c, 2 Hz + 1 Hz (1,000 ms). Theory: The 11 different curves represent predictions as α (the proportion of third-order motion processing vs. first-order) varies from 0 (pure first-order motion, top) to 1 (pure third-order motion, bottom). See the online article for the color version of this figure. AQ: 71

Figure 13 illustrates the outputs Θ that the contrast-strength-vector summation theory with the default parameters ($\rho = 0.49$, $\beta = 2$) would have for plaids with 11 values of alpha between 0 (pure first-order motion) and 1 (pure third-order motion). Consider, first, the high-temporal-frequency data that stimulate only the first-order system, the top curve in Figure 13a, which consists of four overlapping curves measured in Experiment 2. The power-law data derived from these curves were fit within measurement error for each subject by the two-parameter theory: Δ, the ratio of 20- to 10-Hz strengths, and β, the exponent of the contrast power function. The data of Experiment 3, in which overall plaid contrast and the angle between the components was varied, were almost perfectly fit with just one parameter, Δ, which was measured a priori in Experiment 2. The data from the Type I plaids of Experiment 1, in which both components have the same spatial and temporal frequency and vary only in contrast, can be fit with zero estimated parameters. However, more precise data fits for the low-temporal-frequency data in Figure 13 would require a more detailed specification of the third-order motion system—and the current experiments were not designed to expose these details.

AQ: 58

Figure 13 illustrates that the default-parameter theory, with zero parameters estimated from the data and an imperfect third-order motion theory, nevertheless captures the essence of same-spatial-frequency plaid data in full ranges of absolute contrasts, relative contrasts, and temporal frequencies.

Conclusions

On the assumption that perceived direction and speed of motion are determined independently, this study focused exclusively on perceived direction. Highly accurate predictions of the perceived motion direction of plaids were possible without any consideration of their actual speed (which is independent of contrast—the main variable under consideration) or of their perceived speed (which was not determined and not needed). When only the first-order motion system is stimulated, perceived plaid direction is determined entirely by vector summation of the contrast-strength vectors that represent individual sine-wave components. Component contrast strength was found to be a power (with an exponent of 2 ± 0.4) of its contrast. Previously proposed theories involving the vector sum of component velocities are replaced by the vector sum of component contrast-strength vectors. To the extent that third-order motion processing was possible at lower temporal frequencies, it influenced perceived plaid motion direction of Type II plaids toward the direction of rigid translation. For component temporal frequencies of 1 and 2 Hz with stimuli in which both plaid components were of equal high contrast, the perceived motion direction was precisely the direction of rigid translation. There was no influence of first-order motion, that is, the perceived motion direction was that of pure third-order motion. For low-temporal-frequency plaids with unequal component contrasts, the judged direction of motion deviated from rigid toward the direction of the higher contrast plaid component in proportion to the power of the difference in component contrasts, a second direction of motion concurrently computed by the third-order motion system. The four vectors—the two component vectors computed by each of the first- and third-order motion systems—sum to produce the predicted output direction. For same-spatial-frequency plaids, using two default parameters and zero parameters estimated from the data, the contrast-strength-vector summation theory predicts the overall pattern of motion direction judgments over the full ranges of temporal frequencies and contrasts of the plaid components.

References

Adelson, E. H., & Bergen, J. R. (1985). Spatio-temporal energy models for the perception of apparent motion. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 2,* 284–299.

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

Amano, K., Edwards, M., Badcock, D. R., & Nishida, S. (2009). Adaptive pooling of visual motion signals by the human visual system revealed with a novel multi-element stimulus. *Journal of Vision, 9*(3), 4. http://dx.doi.org/10.1167/9.3.4

Anstis, S. M. (1970). Phi movement as a subtraction process. *Vision Research, 10,* 1411–1430.

Appelbaum, L. G., Lu, Z.-L., & Sperling, G. (2007) Contrast amplification in global texture orientation discrimination. *Journal of Vision, 7*(10), 13.

Barlow, H. B., & Foldiak, P. (1989). Adaptation and decorrelation in the cortex. In R. Durbin, C. Miall, & G. Mitchison (Eds.), *Computing neuron* (pp. 54–72). Boston, MA: Addison Wesley.

Blakemore, C., & Campbell, F. W. (1969). On the existence of neurons in the human visual system selectively sensitive to the orientation and size of retinal images. *Journal of Physiology, 203,* 237–260.

Blaser, E., Sperling, G., & Lu, Z.-L. (1999). Measuring the amplification of attention. *Proceedings of the National Academy of Sciences of the United States of American, 96,* 11681–11686.

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

Bowns, L. (2018). Motion estimation: A biologically inspired model. *Vision Research, 150,* 44–54.

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

Burge, J., & Geisler, W. S. (2015). Optimal speed estimation in natural image movies predicts human performance. *Nature Communications, 6,* 7900. http://dx.doi.org/10.1038/ncomms8900

Burr, D., & Thompson, P. (2011). Motion psychophysics: 1985–2010. *Vision Research, 51,* 1431–1456. http://dx.doi.org/10.1016/j.visres.2011.02.008

Carandini, M., Heeger, D. J., & Movshon, J. A. (1997). Linearity and normalization in simple cells of the macaque primary visual cortex. *Journal of Neuroscience, 17,* 8621–8644.

Champion, R. A., Hammett, S. T., & Thompson, P. G. (2007). Perceived direction of plaid motion is not predicted by component speeds. *Vision Research, 47,* 375–383.

Chubb, C., Sperling, G., & Solomon, J. A. (1989). Texture interactions determine perceived contrast. *Proceedings of the National Academy of Sciences of the United States of America, 86,* 9631–9635.

Cropper, S. J., & Badcock, D. R. (2008). Perception of direction of motion reflects the early integration of first and second-order stimulus spatial properties. *Journal of Vision, 8*(4), 12.

Emrich, S. M., Riggall, A. C., LaRocque, J. J., & Postle, B. R. (2013). Distributed patterns of activity in sensory cortex reflect the precision of multiple items maintained in visual short-term memory. *Journal of Neuroscience, 30,* 273–287.

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

Fleet, D. J., Black, M. J., Yacoob, Y., & Jepson, A. D. (2000). Design and use of linear models for image motion analysis. *International Journal of Computer Vision, 36,* 171–193.

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

Kelly, D. H. (1979). Motion and vision. II. Stabilized spatio-temporal threshold surface. *Journal of the Optical Society of America, 69,* 1340–1349.

Lettvin, J., Maturana, H. R., McCulloch, W. S., & Pitts, W. H. (1959). What the frog's eye tells the frog's brain. *Proceedings of the IRE, 47,* 1940–1951.

AQ: 60 Liu, D., & Sperling, G. (2008). The perceived motion direction of fast-moving Type-II plaids. *Journal of Vision, 8*(6), 19.

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

Lu, Z. L., & Sperling, G. (1995b). Attention-generated apparent motion. *Nature, 377,* 237–239.

Lu, Z. L., & Sperling, G. (1996). Contrast gain control in first- and second-order motion perception. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision, 13,* 2305–2318.

Lu, Z. L., & Sperling, G. (2001). Three systems theory of human visual motion perception: Review and update. *Journal of the Optical Society of America*. A, Optics, Image Science, and Vision, *18,* 2331–2370.

Mach, E. (1886/1959). The analysis of sensations and the relation of the physical to the psychical (C. M. Williams and S. Waterlow, Trans.), New York, NY: Dover Publications.

Maruya, K., Amano, K., & Nishida, S. (2010). Conditional spatial-frequency selective pooling of one-dimensional motion signals into global two-dimensional motion. *Vision Research, 50,* 1054–1064.

Mather, G., Pavan, A., Bellacosa Marotti, R., Campana, G., & Casco, C. (2013). Interactions between motion and form processing in the human visual system. *Frontiers in Computational Neuroscience, 7,* 65. http://dx.doi.org/10.3389/fncom.2013.00065

Movshon, J. A., Adelson, E. H., Gizzi, M. S., & Newsome, W. T. (1985). The analysis of moving visual patterns. In C. Chagas, R. Gattass, & C. Gross (Eds.), *Pattern recognition mechanisms* (pp. 117–151). Rome, Italy: Vatican Press.

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

Perrone, J. (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.

Perrone, J. A., & Thiele, A. (2001). Measuring the visual speed analyzing properties of primate MT neurons. *Nature Neuroscience, 4,* 526–532.

Poulton, E. C. (1967). Population norms of top sensory magnitudes and S. S. Stevens' exponents. *Perception and Psychophysics, 2,* 312–316.

AQ: 63 Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 303–317). New York, NY: Wiley.

Sperling, G. (1989). Three stages and two systems of visual processing. *Spatial Vision, 4*(2–3), 183–207.

Sperling, G. (1992). *Type 1 and Type 2 experiments*. Retrieved from http://www.cogsci.uci.edu/whipl/Type_1_and_Type_2_Expts.pdf AQ: 64

Sperling, G., Dosher, B. D., & Landy, M. S. (1990). How to study the kinetic depth effect experimentally. *Perception, 38*(Suppl.), 57. AQ: 65

Sperling, G., & Liu, D. T. (2009). The vector sum of motion strengths describes the perceived motion direction of first-order plaids. *Perception, 38* (Suppl.), *57*. AQ: 66

Stevens, S. S. (1957). On the psychophysical law. *Psychological Review, 64,* 153–181.

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

Sun, P., Chubb, C., & Sperling, G. (2014). A moving-barber-pole illusion. *Journal of Vision, 14*(5), 1. http://dx.doi.org/10.1167/14.5.1

Sun, P., Chubb, C., & Sperling, G. (2015). Two mechanisms that determine the barber-pole illusion. *Vision Research, 111*(Part A), 43–54.

Tomassini, A., Morgan, M. J., & Solomon, J. A. (2010). Orientation uncertainty reduces perceived obliquity. *Vision Research, 50,* 541–547.

Tseng, C.-H., Gobell, J., & Sperling, G. (2004). Long-lasting sensitization to a given colour after visual search. *Nature, 428,* 657–660.

van Maanen, L., Grasman, R. P. P. P., Forstmann, B. U., Keuken, M. C., Brown, S. D., & Wagenmakers, E. J. (2012). Similarity and number of alternatives in the random-dot motion paradigm. *Attention, Perception, and Psychophysics, 74,* 739–753.

van Santen, J. P. H., & Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America. A, Optics, Image Science, and Vision*, *1,* 451–473.

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

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

Wei, X. Y., & Stocker, A. A. (2015). A Bayesian observer model constrained by efficient coding can explain "anti-Bayesian" percepts. *Nature Neuroscience, 18,* 1509–1517.

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience, 5,* 598–604.

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

Wilson, H. R., & Kim, J. (1994). Perceived motion in the vector-sum direction. *Vision Research, 34,* 1835–1842.

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

Received March 20, 2019
Revision received August 7, 2019
Accepted November 5, 2019 ■