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

George Sperling, Peng Sun, Dantian Liu, and Ling Lin

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

George Sperling       sperling@uci.edu
Peng Sun               peng.sun@uci.edu
Dantian Liu            dantingliu@gmail.com
Ling Lin               lincontact@gmail.com
Theory of the Perceived Motion Direction of Equal-Spatial-Frequency Plaid Stimuli

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

Abstract

At an early stage, three 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 two 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 two of the three motion systems are active in plaid processing, and that plaids with temporal frequencies $\geq 10\text{Hz}$ 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 two 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 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 counter-intuitive hypothesis that motion direction is
processed independently of speed at early stages of visual processing.
Introduction

Type 1 and Type 2 Plaids

Because sine waves are basis functions in linear systems analysis, a great deal of research on visual motion perception has been focused on the perception of moving sinewave gratings (e.g., Kelly, 1979, Burr & Thompson, 2011). However, even though real-world stimuli can be considered as being composed of sine waves, pure sinewave 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 sinewave gratings moving in independent directions has been called a plaid (Adelson &Movshon, 1982). Interestingly, every plaid has an interpretation as a single translating frame, i.e., 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 & Movshon (1982) introduced plaids, there has been an enormous concern with the perceptions produced by plaids, >650 publications not including abstracts according to Google Scholar, Feb. 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: (1) Contrary to intuition, motion direction and motion speed are computed separately by the visual system, here we concentrate exclusively on motion direction, and (2) 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.

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 nor various subtleties in the data, but they clearly capture the essence of same-spatial-frequency plaid motion perception.

Adelson & Movshon (1982) observed the motion of two kinds of plaids that Wilson, Ferrera, & Yo (1992) divided into two classes: Type 1 and Type 2 (Fig. 1). In Type 1 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 (Fig. 1). In Type 2 plaids, the direction of rigid movement lies outside of the angle formed by the two component velocity vectors (Fig. 1), making Type 2 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 sinewave gratings. Here, we vary the contrasts of the components of Type 1 and Type 2 plaids and use temporal frequencies above 10 Hz to exclude the third-order motion perception mechanism (Lu & Sperling, 1995a). In this restricted domain, where only the first-order motion mechanism 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.
Figure 1: The Type 1 and Type 2 plaid stimuli and their components used in these experiments. Top row: Type 1 Plaid. Top left: A single frame (snapshot) of Component 1, a 1 cycle per degree (cpd) sinewave with 30% contrast modulation around the mean background level that moves upward to the right within a Gaussian window ($\sigma = 2.0$ degrees of visual angle). The temporal frequency was 10.6 Hz resulting in a speed of $(10.6 \text{ Hz})/(1 \text{ cpd}) = 10.6 \text{ deg/sec}$. 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 deg diameter with a central fixation spot intended to control fixation, vergence, and accommodation. The ticks on the circle help the trained subjects to indicate the direction of perceived movement in degrees (0,...,359). The direction of rigid translation (aka rigid direction, direction of pattern motion, intersection of constraints) is represented by the dashed arrow in the Type 1 diagram. The rigid direction was randomly varied from trial to trial between 0,...,359 deg. The labels C10 and C20, respectively, indicate Components with temporal frequencies of 10.6 Hz and 21.2 Hz, respectively. Middle row: Type 2 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 2 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 Supplementary Information for a video of the above plaids.
Figure 2: Rigid motion direction defined: All motions of sinewave gratings and of plaids (pairs of sinewave gratings) can be produced by viewing a moving snapshot of the grating or plaid through an aperture. (a) A sinewave 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 sinewave grating is inherently ambiguous. (b) The Type 2 plaid used in the experiments. When the two moving sinewave 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 and the vector-sum-of velocities direction. Although a brief view of a moving plaid is logically sufficient to define the rigid direction, i.e., the direction in which the snapshot of the plaid is moving, the rigid direction usually is not the perceived direction of motion.
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 sinewave grating. Consider a snapshot of a sinewave 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. 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 Fig. 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 and Movshon, 1982, Movshon, Adelson, Gizzi, & Newsome, 1986), the ambiguity is resolved. Any plaid—indepedent 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, i.e., 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., Fig. 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 Fig. 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, e.g., luminance, is a function of $x,y,t$. The motion system produces as its output a time-varying vector flow field, i.e., 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 flowfield.

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 that are perceived as figure, small for points interpreted as background. The perceived motion direction of plaids has been interpreted both in terms of factors (such as velocity) that directly affect perceived motion direction and in terms of the contribution of different motion systems to perceived direction.

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, & Muligan., 1990; Weiss, Simoncelli,
Several alternative computational theories have been proposed. In the Vector Summation of Velocities algorithm (VS), 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.

A feature tracking explanation of plaid motion perception was proposed initially by Adelson & Movshon (1982) 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 (Lu & Sperling, 1995b; Blaser, Sperling, & Lu, 1999; Tseng, Gobell, & Sperling, 2004); the third-order motion system computes motion on a spatio-temporal salience field in the same way that the first-order system computes motion on a spatio-temporal luminance field (although more correctly a point-contrast field as the input to the first-order system is interpreted as having both positive and negative values, whereas luminance is always positive). 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 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 1.0 Hz will be used.

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.

Outline

Four experiments are conducted to investigate the perceived motion-direction of plaid stimuli that have sufficiently high-temporal frequencies that they are expected to stimulate only the first-order motion-perception system. All plaid sinewave components are 1 cycle per degree (cpd). Experiment 1 compares motion-direction judgments in two paradigms to determine which is more appropriate to use for Experiments 2-4. Experiment 2, the main experiment, investigates the perceived direction of Type 2 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 are incorporated in a ”Contrast-Strength-Vector-Sum theory” that is used to predict the perceived direction of new plaid stimuli. The model is tested in Experiment 3 which investigates perceived direction of plaids composed of the same 10 and 20 Hz and 1 cpd frequency components as in Experiment 2 except that the the angle between the two components is varied over the full range. Finally, a pure theory, with zero parameters estimated from the data, is shown to capture the essential features of the perceived direction of same-spatial frequency plaids.
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, e.g. 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 zero 0 to 359 deg. Subjects were trained, with feedback, to directly estimate motion directions of moving sinewave 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, Dosher & Landy, 1990; Sperling, 1992). Before each new block of trials, subjects were shown sample stimuli to familiarize themselves with the stimuli they were to judge in that session.

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 coincides with the experimenters’ observations, these additional response options were discontinued.

Stimuli All stimuli in these experiments were composed of sinewave gratings (sinusoidal modulations around the mean luminance); two such gratings were added together to form a plaid (Figs. 1 and 2). The spatial frequency of all gratings was 1 cycle per degree of visual angle (cpd). 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 4-5 dva, as illustrated in Fig. 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, 0,...,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**

Using our method of direct direction estimation, we compared two paradigms: (1) *Random directions* in which the direction of rigid motion on each trial is chosen randomly from the whole circle (0, 359 deg) and (2) *restricted directions* in which the direction of rigid motion is chosen mainly from close-to-vertical directions (-4, 4 deg). Random direction procedures typically are used in more cognitive tasks (e.g., Emrich, Riggall, LaRocque & Postle, 2013) whereas restricted directions typically are 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., Maanen, Grasman, Forstmann, Keuken, Brown & Wagenmakers, E. J., 2012), the two psychophysical methods have not been compared to determine their differential suitability for judging motion direction (our task).

**Stimuli.** Symmetrical Type 1 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 degrees 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%, 0%). The 90 deg angle between gratings causes the second-order motion system
to be ambiguous (see below, section “Second-order motion contributions to plaid motion” ),
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 unequal contrasts, one plaid had higher
contrast counterclockwise to the rigid motion direction of the plaid, the other, clockwise.
For pairs of equal contrast i.e. (2%, 2%), the same plaid was used twice. Figure 3 shows
snapshots of the plaids (rigid direction vertical) in one of their possible orientations.

Figure 4, top row, shows vector diagrams of the two component 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 by 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 divided the whole
luminance range into 256 gray levels (Lu & Sperling, 2001). The mean luminance of the
monitor was 118.6 cd/m². Subjects viewed the display binocularly at a distance of 50 cm in
a darkened room.
Subjects. Two subjects aged 20, 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 IRB of the University of California, Irvine prior to their participation. Subjects were paid for participation.

Experiment 1 procedures. 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 static screen with the plaid stimulus removed, i.e. the vergence circle, the peripheral marker and the central fixation spot remained on the screen at all times. 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 sinewave stimuli that were interleaved throughout the experiments were similar to their prior training stimuli.

Sessions lasted approximately an hour. Two experimental conditions were conducted separately in each experimental session: random directions (0, 359) deg and restricted directions (-4,+4) deg. In the random-directions (0, 359) deg condition, the whole 360 deg circle was
equally divided into 18 sectors, with each sector spanning 20 degrees. Then a direction (in integer degrees) was randomly chosen from a uniform distribution within each 20 deg sector. In the restricted-directions (-4,+4) deg condition, each of the nine close-to-vertical directions of motion (-4, -3, -2, -1, 0, 1, 2, 3, 4 deg) was used twice. A total of 36 directions (18 for the random-direction 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 pseudo-random 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 zero degrees by subtracting the rigid direction from each observation. Given the symmetrical nature of Type 1 plaids, results from different-contrast mirror-image plaids were pooled after flipping one of the plaids around the rigid direction (zero deg). Results for plaids with equal contrast components were simply pooled together. Thereafter, 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 empirical probability distribution plot. The mean and standard deviation of the fitted Gaussian distribution were used to summarize the data.

Results
The left and right halves of Fig. 4 show the data obtained in the two different range-of-motion-direction protocols, (0, 359) deg and (-4, +4) deg, 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 deg and that of the higher-contrast component is +45 deg.

The five curves in Fig. 4 represent individual plaid contrast ratios illustrated in Fig. 3. It is immediately evident that in both protocols both subjects judge equal contrast plaids (2%,
Figure 4: Experiment 1. Cumulative histograms of the judged directions of plaids in which the direction of rigid motion varies randomly between (0, 359) deg or between (-4, +4) deg. Thumbnail images represent 3 of the 5 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, 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 (0,359) deg but not in (-4,+4) deg paradigm. (A,B): Data from subject FT. (C,D): Data from subject ROJ. In each panel, the abscissa indicates the direction of motion $\theta$ relative to the rigid direction; the ordinate indicates the cumulative probability that a judged direction is greater than $\theta$. 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 upwards vertical. See text for details. Single-component plaid directions (pure sine waves) are judged correctly for the (0, 359) deg conditions (A and C) but incorrectly for restricted directions (-4, +4) deg conditions (B and D). The two lowest contrast components have a very small influence on perceived direction.
2% as moving in the rigid direction (0 deg) 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, i.e., 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) deg conditions (Fig. 4A,C), but are judged incorrectly for (-4, +4) deg conditions (Figs. 4B, 4D). In particular, for the (0, 359 deg) condition, the judged single-component direction was 46.0 degrees for subject FT (Fig. 4A), and 45.5 deg for subject ROJ (Fig. 4B), both judgments are within measurement error of 45 deg. However, for the (4, +4 deg) restricted directions condition, large deviations from the veridical 45 deg direction were observed: 64.2 deg for subject FT (Fig. 4B) and 52.3 deg for subject ROJ (Fig. 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 towards cardinal orientations (attraction) depending on noise type and its magnitude (Tomassini, Morgan & Solomon, 2010; Wei & Stocker, 2014). In those tasks, systematic biases are never found for oblique orientations ±45 deg; the Bayesian explanations of those biases also predict zero bias for ±45 deg orientations. The very large contextual bias in ±45 deg 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) deg 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 the (0,359) deg 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) deg random-directions paradigm.

The large effect of contrast on perceived direction. Van Santen & Sperling (1984) showed that the Elaborated Reichardt Model accounted astonishingly well for five counter-intuitive predictions of phenomena of first-order sinewave motion. The Motion Energy Model (Adelson & Bergen, 1985) and Hilbert transform Model (Watson & Ahumada) 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 sinewave. 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 the 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) deg 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) deg random-directions paradigm. Before considering the theoretical implications, more data are needed.

Experiment 2: Perceived motion direction of Type 2 plaids as function of component contrast and temporal frequency

Experiment 2 is an extensive empirical study of mostly high-temporal frequency Type 2 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 1 plaid stimulus from Experiment 1, and a Type 2
plaid stimulus from Experiment 2. The diagrams in Fig. 1 make obvious the advantages, originally noted by Ferrera and Wilson (1990), of Type 2 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 1 plaid, as in Fig. 1, the directions of rigid translation and of the vector sum of component velocities are identical. For the Type 2 plaid, the directions of rigid translation and vector sum of velocities are very different.

**Stimuli**

Type 2 plaids were composed of two components, both sinewave gratings, here designated 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 deg clockwise relative to the rigid direction; C10 moved 70.5 deg clockwise relative to the rigid direction (Fig. 1). Both components had a spatial frequency of 1.0 cpd. C20 translated 90 deg in phase between consecutive frames; C10 translated 45 deg in phase between consecutive frames. Stimulus duration was 200 ms, i.e., a total of 17 stimulus frames at a frame rate of 85Hz.

A total of 9 different component contrast ratios (Fig. 5a) was explored times 4 conditions with different maximum contrasts. Contrast ratio is “contrast of lower temporal frequency component/contrast of higher temporal frequency component”, which was simply contrast(C10)/contrast(C20). The contrast ratio took values of 0, 1/8, 1/4, 1/2, 1/1/, 2, 4, 8, inf. 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 5 stimuli of Fig. 5a.
Procedure.

The data were collected in two phases, which differed slightly in the way in which pseudo-random directions were chosen. In phase 1, a circle was equally divided into 36 intervals of 10 deg. A direction (in integer degrees) was randomly chosen from a uniform distribution within each 10 deg interval thereby producing 36 different directions. In phase 2, a direction was chosen from each 4 degree interval around the circle, yielding 90 different directions. Trials were run in a pseudo-random mixed-list design, i.e., 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 and 15 Hz replacing the original 21.2 and 10.6 Hz.

Subsequent to the main experiment, two separate sessions were run which used only the highest-contrast stimuli in which one component’s contrast was always 32%: (1) The temporal frequencies of the plaid components were 10.6 Hz and 5.3 Hz (designated as 10:5 Hz) with a stimulus duration of 765 msec; and (2) temporal frequencies of 2.02 Hz and 1.01 Hz (2:1 Hz) with a stimulus duration of 1000 msec.

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 10.3 % data (out of a total of 126 data points) were excluded from analysis, and 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 2 plaids as a function of the contrasts and temporal frequencies of the components. (a) The 32% contrast set of 9 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 \( \theta \) relative to the rigid direction. The angles of the two component gratings (48.2 deg, 21.2 Hz; 70.5 deg, 10.6 Hz) and the rigid direction (0.0 deg) 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 \( \theta \). 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 9 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 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-f are 10 and 20 Hz (pure first-order stimuli), the plaid with equal-contrast components of 1 and 2 Hz (panel h) stimulates pure third-order motion precisely in the rigid direction. See Supplementary Information for video examples of the above plaids.
Figure 6: Judged directions of randomly oriented Type 2 plaids as a function of the contrast ratio and temporal frequencies of the components, data for 3 subjects. The three thumbnail images above each panel illustrate the two pure sinewave 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, 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 depends on relative contrast, all are indicated as horizontal lines. At temporal frequencies sufficiently high to bypass the third-order system (20Hz,10Hz panels a,b; 30Hz,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 Supplementary Information for video examples of the above plaids.
Results and Discussion.

Figure 5b shows cumulative probability distribution functions (as in Fig. 4) of motion direction responses for subject ROJ viewing Type 2 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 deg. These PSE are within a degree of 48.2 deg, 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 sinewave grating. Alternatively stated, for plaids with contrasts (32%,8%), (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 deg which, at this scale, is hard to distinguish from 55.6 deg, the Vector Sum (VS) of the component velocities. However, when both gratings are of equal contrast (32%,32%), the PSE is 64.7 deg. Equal-component-contrast plaid motion is judged much closer to the 10 Hz component direction than would be expected from VS.

In the panels b-h of Fig.5, only the PSEs are represented, not the probability distributions from which they were derived. The PSEs in Fig.5a re-appear in the left half of Fig.5f. The Standard Error of the PSEs in Fig. 5 averages less than 0.5 deg for all three subjects, i.e., the 95% confidence interval is less than ±1 deg 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 and 10Hz, and the display duration increased to 765 msec, the judgments of approximately equal-contrast component plaids approach the rigid direction. At component frequencies of 1 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-f are remarkably similar. To better display the relations between the 6 data curves in Fig. 5 c-h, they are all plotted in...
the same graph in Fig. 6a. Figure 6 also shows data from two other subjects. An astounding result in Fig. 6 is that when the component temporal frequencies are 10 Hz and 20 Hz for subjects ROJ, AL, and 15 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%), (4%, 2%) are statistically identical. That judged direction depends only on contrast ratio (not on the actual contrasts) obtains for all other contrast ratios, from zero to infinity.

A second significant property of the data in Fig. 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 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 towards the rigid direction, no trace of an IOC computation.

The 20:10 Hz data for subject FT in Fig. 5c differ from those of 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 towards the rigid direction. Perceived directions of plaids with contrasts of 32% and 16% deviate so much towards the rigid direction that the perceived direction no longer lies between the directions of the components.

We interpret the data of subject FT in Fig.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 and 10 Hz trials, except that the plaid component frequencies were now increased to 30 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 plaid's 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 plaid's with contrast ratios of 1/2 and 1 (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 Fig. 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 Fig. 5) is treated in detail in Experiment 3. The
data of Experiment 2 give no indication that second-order motion plays any role in the
perception of the motion direction of these plaid stimuli.

**Contrast-Strength-Vector-Sum theory for Motion-Direction Perception**

**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-Sum 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, depth,
Figure 7: The Contrast-Strength-Vector-Sum theory and the resulting power law: Contrast strength is a power function of stimulus contrast. (a) Estimating the relative contrast strength $S$ of the two components of a plaid from the plaid’s judged direction. Plaid components are represented as vectors that have a direction equal to the direction of the component sinewave and a length proportional to its contrast. Both plaid components have a spatial frequency of 1 cpd and a contrast of 32%. Component 1, $C_{10,32}$, has a temporal frequency of 10 Hz and moves directly upward; Component 2, $C_{20,32}$, has a temporal frequency of 20 Hz and moves at an angle of $\phi_{22}=22.3$ deg relative to component 1 (as in the Experiment 2). If the components were of equal strength $|C_{32,10}| = |C_{32,20}|$, then the vector sum $|C_{32,20}| + |C_{32,10}|$ would lie on a direction $\phi_{22}/2$ exactly between the two components. However, the judged direction $\phi_j < \phi_{22}/2$. The geometric construction in (a) shows that, if the contrast of the 20 Hz component were reduced by a factor $\rho < 1$, the vector sum $\rho |C_{20,32}| + |C_{10,32}|$ would indeed match the judged direction. Component 2, $C_{32,20}$, is a vector, its contrast-strength $S_{32,20} = \rho |C_{20,32}|$ is a scalar. The strength of the strongest stimulus in the experiment $|C_{32,32}|$ is arbitrarily defined as 1.0. The geometric construction enables estimation of the relative strength $\rho$ of any two components, and the estimation of the absolute strength of component 2 whenever the absolute strength of component 1 is known. (b) Contrast strengths of the Type 2 plaid stimuli of Experiment 2 derived from motion-direction judgments by the Contrast-Strength-Vector-Sum theory. The ordinate represents the contraststrengths of the gratings components 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 gratings (subjects ROJ, AL), 15 Hz (FT). Unfilled dots indicate 20 Hz gratings (subjects ROJ, AL), 30 Hz (FT). The straight lines are least squares best fits to the data; they are constrained to have the same slope for all data from a subject; they represent power laws with exponents of approximately 2 (see text for details); and they account for 99% of the data variance. Only plaid components with estimated contrast strengths that differed significantly from zero-strength are shown and incorporated into the data fits. The brackets on top indicate the four ranges of stimuli within which useful strength comparisons were possible, they correspond to data within the four panels c,d,e,f of Fig. 5.
seem to be computed in different specialized brain areas and within these computational areas, different scales (image resolutions) are treated separately. So, it is not unreasonable to assume that motion speed and direction may be computed separately in the early stages of visual processing. (See subsection on Neural economy, below.) The computation of exclusive concern here is perceived motion direction.

The vector diagram in Fig. 7a illustrates how the contrast vector strength of a plaid component grating is measured. It is assumed that each sinewave component of a plaid stimulus can be represented as vector that has a direction—that of the sine wave motion—and a magnitude—determined by contrast and temporal frequency. We begin by assuming that the strongest motion component in these experiments, 10 Hz, 32% contrast, has a contrast strength of 1.0. Figure 7a shows a particular example from Experiment 2 to illustrate the Contrast-Strength-Vector-Sum theory—a plaid composed of two 1 cpd sinewave gratings, one 10 Hz 32% contrast and the other 20 Hz 16% contrast, with an angle of 70.5 - 48.2 = 22.3 deg between the components. 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. The question is: Given the contrast strength of 1.0 for the 10 Hz 32% contrast component, what length of vector in the direction of the 20 Hz 16% grating would produce the perceived direction. That length, which is less than 1 (by a factor of $\rho$ in Fig. 7a), is the contrast strength $\rho|S_{16,20}|$ of the 20 Hz 16% grating component relative to 1.0.

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, i.e., the 10 Hz 32% contrast sine wave, the relative strengths become absolute strengths (relative to the standard). Using a simple algebraic formulation of the estimation procedure described above, the contrast-strengths of all the motion components in Experiments 1 and 2 were estimated. Because the same motion component may occur in up to 4 different plaids in Experiment 2 (and more in Experiment 1), 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 Fig. 7b. It shows the contrast strengths of the sinewave components of plaid as a function of their contrasts with temporal frequency as a parameter. Only plaid-component sinewaves that were processed exclusively by the first-order motion system (10 and 20 Hz for subjects AL, ROJ, and of 15 and 30 Hz for subject FT) and only plaid components that were significantly different from zero are included in the contrast-strength estimates. The straight lines that are fit to the contrast strengths of component sinewaves are constrained to have the same slope for both of each subject’s component temporal frequencies. On a log-log, plot straight lines represent power laws (contrast strength increases as power of contrast) with exponents approximately equal to 2. For subjects FT, AL, ROJ, the slope exponents are (1.6, 2.0, 2.4). The difference between the curves is $\rho = (\text{strength of } 20 \text{ Hz})/(\text{strength of } 10 \text{ Hz}) = (0.45, 0.53, 0.33)$ for the three subjects. 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. Fig. 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 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 power-law neural effect 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 feedforward gain control mechanism provides a much better explanation (Sperling, 1989). How a functional power law is represented in the brain (and vice versa) has not been determined. A power law representation of stimulus inten-
sity 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 threshold
increment decreases as the size of the base increment increases above zero before the incre-
ment threshold inevitably increases towards a Weber fraction. The wide range of precise
power law description of behavior in Fig. 7 is the result of 4 overlapping ranges—not of a
wide range from within a single condition. We are not aware of any other similarly precise,
wide-range, power law description of an unconscious perceptual computation.

All high temporal frequency (first-order) stimuli were presented in a mixed list. So 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 msec during
which the stimulus was presented. How this can happen is itself a very provocative question.

Beyond deriving a power law, to be useful, the Contrast-Strength-Vector-Sum theory
must apply in other contexts. We demonstrate this in Experiment 3, but to understand
Experiment 3, we first need to define explicitly 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), e.g. 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 flowfield
have a direction and a (non-negative) 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 flowfields 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
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) 1st-order: The center disk has equal physical contrast in (a) and (b). It appears darker in (b) than (a) because in (b) it is surrounded by a higher mean luminance. (c,d) Second-order. The center disk in (c) and (d) has the same physical texture contrast. It appears to be of lower contrast in (d) than (c) because in (d) it is surrounded by a greater texture contrast (higher variance texture). (e) First-order motion. The sine wave in each of the 5 frames in (e) differs from the adjacent frame by 90 deg. When the five frames are presented successively (from top to bottom), superimposed, (e) represents a “luminance” motion stimulus moving to the right directed to the first-order motion system. Alternatively, (e) as illustrated represents a first-order sinewave grating slanted to the right. (f) Second-order motion. Each of the frames in (f) represents a sinewave modulation of texture contrast, The contrast modulation in each of the 5 frames in (f) differs from the adjacent frame by 90 deg. Presented successively (from top to bottom), superimposed, (f) represents a ”contrast” motion stimulus moving to the right directed to the second-order motion system. Alternatively, (f) as illustrated represents a second-order sinewave grating slanted to the right. (g-j) Third-order motion. Four frames in which the central rectangle (“figure”) is differentiated from it’s surround (“ground”) by, successively, (g) slant, (h) contrast, (i) stereo-depth, (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. These motion examples are merely illustrative, actual experimental stimuli are more complex. [a-d after Figure 1, p. 9632 in Chubb et al. (1989), Copyright (1989) National Academy of Sciences, USA, by permission; e,f after Figure 7, p. 2338, and g-j after Figure 4, p. 2335, in Lu & Sperling, 2001, Optical Society of America, by permission.]
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.

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 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 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 compli-
cated than merely computing Weber-contrast (i.e., luminance adaptation): it involves spatial
filtering into spatial frequency channels and orientation filtering within channels (Blakemore
and Campbell,1969), and light adaptation and contrast-gain control (Lu & Sperling, 1996),
all before the motion computation.

The computations in second-order motion (as defined by Lu and Sperling, 1995, 2001)
involve all the processes listed for first-order motion, the main difference being 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, i.e. 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, 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 pre-processing prior to third-order motion (as defined by Lu and Sperling, 1995, 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, where "figure" areas have higher values of salience than "ground" areas. Whereas figure-ground is often thought of as a binary variable, salience is 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 (Lu & Sperling, 1995b; Blaser, Sperling, & Lu, 1999; Tseng, Gobell, & Sperling, 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 have not yet been worked out. What is clear, however, is that third-order motion is inherently concerned with the motion of objects, areas of salience. The first-order system does not know about objects, it merely computes local flowfields. Relating first-order motion flowfields 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, i.e, the input to second-order motion processing is a function of the absolute
Figure 9: Motion direction of a Type 2 plaid according to a second-order motion computation. (a) Snapshot of a Type 2 plaid. (b) The same stimulus as (a) except that the point-contrast values have been squared (rectified). The two solid lines in (b) indicate the positive peaks of the fundamental sinewave—the dominant sinewave component—in the rectified stimulus pattern. The two dotted lines indicate the negative peaks of the dominant sinewave component. The arrow indicates the motion direction of the dominant second-order sinewave. (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.

value of the local contrast. This function can be simply 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 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

Figure 9a illustrates a snapshot of a plaid stimulus; Fig. 9b, the squared contrast of the stimulus; Fig. 9c, a vector diagram illustrating the direction of motion of the dominant sinewave component of the squared stimulus as well as the directions of the original sinewave components. (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 2 plaid stimulus from Expt. 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

¹The 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 (Fig. 2).
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 sinewave 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 case of equal contrast, 1 cycle/deg component gratings, 10 and 20 Hz, 22.3 deg between the components, the amplitude of the strongest second-order component is 0.50 times the amplitude of the the original stimulus components. The second-order component in the stimuli of Experiment 2 has a large stimulus amplitude. In Fig. 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-Sum 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, possibly different temporal frequencies \( f_1, f_2 \), contrasts \( c_1, c_2 \), and strengths \( S_1(f_1, c_1), S_2(f_2, c_2) \). Assume that the relative strength \( \rho = S_2/S_1 \) is known, e.g., 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 \( S_1 \) and \( S_2 \) stimulate only the first-order motion system, \( \rho \) is sufficient to enable prediction of the perceived direction of any plaid composed of components \( S_1 \) and \( S_2 \) no matter what angle between them may be. For
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 5 continuous curves representing theoretical predictions of 4 theories. As in Experiment 2, one grating component of the plaid has a temporal frequency of 20 Hz, the other 10 Hz; the contrasts of the data points are: 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 constant $\rho$ from Experiment 2 (for subjects FT and ROJ) and a $\rho$ estimated from only the 20 deg 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 are the predictions for Vector Sum of Velocities, Rigid direction, and the direction of Second-order motion. The direction of Second-order motion flips 90 deg as the angle between the gratings changes from 89 to 91 deg; the strength of Second-order motion approaches zero as the angle between the gratings approaches 0 or 180 deg. The solid straight lines (left panel only) show the representation in this graph of the directions of the plaid components $C_{20}$, $C_{10}$. 
any new plaid, the direction of the sum of the components’ contrast-strength vectors predicts
the perceived motion direction, and $\rho$ is sufficient to determine the relative strengths.

To test the Contrast-Strength-Vector-Sum theory, Experiment 3 takes two different pairs
of gratings from Experiment 2 whose relative contrast strength has been determined in
a Type 2 plaid with an angle of 22.3 deg 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-Sum
theory is correct, the outcome of Experiment 3 will be perfectly predicted from the 22.3 deg
observations in Experiment 2, with zero parameters to be estimated in Experiment 3.

Procedures

In Experiment 2, the angle between the two components of the plaid was 22.3 deg, contrast
and temporal frequency were varied. Here, six angles between 20 to 120 deg are used (0 and
180 deg are not useful) and the spatio-temporal frequencies of the components are the same
as those used in Experiment 2, 1 cpd, 10 and 20 Hz, . Two pairs of contrasts are used: In
both pairs, the 20 Hz grating had a contrast of 4%, in one plaid the contrast of the 10 Hz
grating also was 4%, in the other, 2%. As in Expt. 2., the rigid direction was randomly
chosen on each trial from 0-359 deg, 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

Figure 10 shows the results (perceived 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 Fig. 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 towards the 10 Hz direction, but remains always in between. The movement of the data away from the 20 Hz direction is greater, obviously, for the stronger 4% contrast 10 Hz component than for the 2% contrast 10 Hz component. The lines through the data points use the contrast-strength ratio $\rho$ derived the data for 22.3 deg plaids in Experiment 2 for the two subjects who participated in Experiment 2 and, for the new subject, LL, from the Experiment 3, 20 deg data. The same $\rho$ that describes the relative contrast strength of two components of a plaid in Expt. 2 also predicts quite accurately how these components will combine in the 7 new plaids of Experiment 3.

In Fig. 10, the predicted curves for both component contrast pairs (4%, 4%) and (4%, 2%) for the two subjects who participated in Experiment 2 account for over 97% of the variance of the data; that is zero-estimated parameters from Experiment 3 itself, i.e., pure prediction. In contrast, Vector Sum of velocity, 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 $\rho$ of the two sinewave 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 plaid (Fig. 11). Orientation judgments cannot be used exclusively to judge motion direction because
Figure 11: Pattern-orientation judgments in a snapshot of a Type 2 plaid. (a) Snapshot of a Type 2 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 to a motion-direction judgment, the motion direction is assumed to be perpendicular to the pattern-orientation. (c) The various judgments and theories compared.

A snapshot of a plaid is same, independent of the temporal frequencies of the components which determine a wide range of different rigid (objective) and perceived motion directions. The influence of orientation *perse* obviously is 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 Fig. 11. The method and results are fully described in Appendix. The main conclusion is that orientation judgments have a much smaller variance than motion judgments and a much smaller variation with intermediate contrast ratios of the components than do motion judgments.
The Contrast-Strength-Vector-Sum theory for the perceived motion direction of equal-frequency plaid stimuli

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, where the physical motion direction of the stimulus was randomly selected from a 0-359 deg range of directions, led to unbiased judgments of perceived motion direction.

The main experiment, Experiment 2, investigated Type 2 plaids composed of two 1 cpd gratings. One component moved in a direction 70.5 deg from the rigid direction, the second component with twice the temporal frequency of the first moved 48.2 deg relative to the rigid direction. For temporal frequencies of 10 and 20 Hz (two subjects) and 15 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 2 plaids (contrast =32, 16, 8, 4 %) was 63±2 deg (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 deg). For smaller contrast ratios, intermediate directions between the two components’ directions were systematically observed.

When component temporal frequencies were decreased 5 and 10 Hz (2 subjects), 10 and 20 Hz (one subject) into temporal frequency ranges where third-order motion was possible, the
perceived direction of plaids with equal or near-equal component contrasts, particularly high absolute contrasts, deviated increasingly towards the rigid direction as temporal frequency was reduced. For our Type 2 plaids with component temporal frequencies of 1 and 2 Hz and contrasts of 32%, all subjects perceived the plaid as moving precisely in the rigid direction, i.e., pure third-order motion with no measurable influence of first-order motion.

Using a simple contrast-strength vector model, the relative contrast strength $\rho$ 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 sinewave 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 $\rho$, 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 deal (or cannot deal) with these basic findings, and we propose a new theory.

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 “Intersection of Constraints (IOC)” algorithm (e.g., Fig. 1). However, for plaids with unequal contrasts, the perceived direction was biased towards the higher-contrast component (Stone et al., 1990). However, the rigid direction is independent of contrast. As a result, Stone, Watson and Mulligan (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 (Campion et al, 2007). Therefore, Champion 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.

Fleet (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.

Vector Summation that combines Fourier and non-Fourier (second-order) component motion signals was proposed by Wilson, Ferrera and Yo (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 using the same second-order computation as the Wilson group.

Weiss, Simoncelli and Adelson (2002) developed an optimal Bayesian estimator which 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 intersection of constraints (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 2 plaids in Experiment 2, for temporal frequencies above 10-15 Hz, the direction of perceived motion always lies between the two components’ directions and never in the rigid direction which was 48 deg 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, Marotti, Campana & Casco, 2013; Sun, Chubb & Sperling, 2015; Sun, Chubb & Sperling, 2014). In particular, Amano et al. (2009) showed subjects arrays of non-overlapping translating Gabor patterns. When each half of all the Gabor patterns translated in a direction with a temporal frequency that was consistent with a Type 2 configuration, two out of three subjects showed a bias towards the rigid direction. This result is interesting because non-overlapping Gabor patterns do not produce features that move in the rigid direction, so any perception in the rigid direction must have been produced otherwise, possibly by an intersection of constraints computation. However, unlike Experiment 2 (above) in which all subjects observed 100% perfect rigid direction perception, Amano et al’s (2009) 2/3 subjects observed only a weak bias towards 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.

Rational for considering only direction, not the speed of perceived plaid motion.

1. Neural economy. 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, i.e., a velocity neuron, then 10x10=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 $10^n$ versus $10^{n+1}$ 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.

2. **Contrast sensitivity of direction-sensitive neurons.** There are three more-or-less equivalent models of early visual motion-direction processing (Adelson & Bergan, 1985; van Santen & Sperling, 1984; Watson & Ahumada, 1985) that all are equivalent for reasonable ranges of their parameters (Sperling & van Santen, 1985), and therefore all account for the 5 counterintuitive results reported by van Santen & Sperling (1984). The motion detector in each of these models (Motion Energy, Reichardt, Hilbert) produces spatially localized vectors that represent a direction and the strength of motion in that direction—where 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 useful for speed detection. However, several more elaborate computational schemes have been proposed for deriving speed from arrays of such detectors, e.g., Heeger (1987), Perrone (2012), and from such arrays enhanced with other types of detectors, e.g., Burge & Geisler, (2015).

3. **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 sinewave 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, Pitts, & McCulloch (1956), 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, 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 stimulus such as plaids...
Figure 12: Flowchart of the Contrast-Strength-Vector-Sum 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 filter attenuates frequencies greater than about 10 Hz. 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, approximately 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 amount by which the larger amplitude exceeds the smaller-amplitude sine wave). In the present experiments, the relative contribution $\alpha$ of the first- and third-order systems depends only on the temporal frequencies of the plaid components. In this illustration, $\alpha=0.5$. The output direction is the $\alpha$-weighted sum of the four vectors computed by the two motion systems.
do not represent the broad-band 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, i.e., dark areas, 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, i.e., speed and direction, is unambiguously and best estimated from the movement of a point or of a feature, i.e., a broadband input that contains a wide range of sinewaves in coherent phase. Insofar as the third-order motion system computes the the direction of rigid translation of Type 2 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.

**Formal specification of the Contrast-Strength-Vector-Sum theory.**

**Four assumptions.** The theory considers only equal-spatial-frequency plaid stimuli. It takes as input the two plaid sinewave 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 ($\theta_1, \theta_2$). This is a large 6-dimensional input space. The theory output is a predicted perceived direction, i.e., 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:

1. The first-order system processes input plaid components of different angles in-
dependently. For each input component, it produces a two-dimensional vector output
\[ C_j((\rho_j c_j)^\beta, \theta_j); \] where \( \rho_j \) is the component’s relative contrast strength \( (\rho \leq 1) \), \( c_j \) is its contrast, \( \theta_j \) is its motion-direction angle, \( \beta \) is 2\( \pm \)0.4, and \( j=1,2 \) designates a component.

(2) The third-order system’s two components are: (2a) a rigid-direction plaid-component \( C_r((2c_n)^2, 0) \) with amplitude \( (2c_n)^2 \) where \( c_n = \min(\rho_1 c_1, \rho_2 c_2) \) and direction 0, and (2b) a sinewave component \( C_s(\|\rho_1 c_1 - \rho_2 c_2\|^2, \theta_m) \) with amplitude \( \|\rho_1 c_1 - \rho_2 c_2\|^2 \) and direction \( \theta_m \), \( c_m = \max(c_1, c_2) \).

(3) The first- and third-order motion system’s outputs are motion-direction vectors \( V_1, V_3 \), each of which is the vector sum of its system’s two component-direction vectors:
\[ (3a) \ V_1(L_{V1}, \theta_{V1}) = \text{VectorSum}(C_1, C_2) \]
\[ (3b) \ V_3(L_{V3}, \theta_{V3}) = \text{VectorSum}(C_r, C_s) \]
where \( L \) and \( \theta \) are length and angle of the output vector; the subscripts 1,2 refer to the plaid components; and the subscripts r and s refer to the perfect plaid (rigid direction) and to the residual sine component.

(4) The net predicted direction \( \Theta \) depends on the frequency-dependent relative strength \( \alpha \) of the first- and third-order system outputs, \( V_1, V_3 \). As the experiments measure only the direction of judged motion, not its magnitude, the theory assumes that the predicted perceived angle \( \Theta \) is simply the \( \alpha \)-weighted mean of the outputs of the first- and third-order systems:
\[ \Theta = (1-\alpha)\theta_{V1} + \alpha \theta_{V3} \]

**Default model.** 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, e.g., the remarkably linear power-law predictions (Fig. 7), and the predictions for plaids with different component angles (Fig. 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 perfect plaid plus residual sinewave nor of the assumed square-law
combination of these two components. The aim here is to demonstrate that these assumptions are nevertheless sufficiently to enable, *a priori*, a prediction of our entire data set with default parameters.

Three parameters that are necessary to predict the data: $\beta$, the power law exponent that determines the relative strength of different values of component contrast; $\rho$, the relative contrast strength of different temporal frequencies; and $\alpha$ which determines the proportions of first- and third-order motion. A basic property of the Reichardt and similar models is that the strength of sinewave motion components is proportional to the square of their amplitude. This was empirically confirmed by van Santen & Sperling (1984). So, the default $\beta$ is 2.0.

Although relative strength $\rho$ of plaid components with different spatial frequencies is easily estimated in experiments, a useful default approximation to $\rho$ is the ratio of detection thresholds—a reciprocal measure of strength. (Recall: The strength $\rho_{10}$ of 10Hz was arbitrarily chosen as 1.) Based on Kelly (1979, Fig. 3, p.1342), the threshold ratio (threshold 10Hz)/threshold 20Hz) =0.244. As thresholds are related to output strength, which is the square of contrast, $\rho_{20} = \sqrt{0.244} = 0.494 \approx 0.5$, which is similar to observed values (0.45, 0.53, 0.33). In the present experiments, $\alpha$ which determines the relative proportions of first- and third-order motion depends on the plaid temporal frequencies.

Figure 13 illustrates the outputs $\Theta$ that the Contrast-Strength-Vector-Sum theory with the default parameters would have in response to plaids with temporal frequency combinations that induce eleven values of $\alpha$ between 0 (pure first order motion) to 1 (pure third-order motion) Precise data fits would require more detailed specification of some of the sub-processes involved particularly third-order motion (the current experiments were not designed to expose all these details) and more parameters to account for individual differences. The comparison between the data and the theory illustrates that this default-parameter theory, with no parameters estimated from the data, captures the essence of the data in full ranges of relative contrasts and of relative angles of the components, and a wide range of temporal frequencies.
Figure 13: Data and theory for one subject judging the motion direction of Type 2 plaid stimuli as a function of the contrast ratio of the component sinewaves (abscissa) with overall contrast and temporal frequency as the parameters. All component sine gratings were 1 cpd. Data are reproduced from Fig. 6a (above). Ordinates: Data: Judged plaid motion direction relative to the direction of rigid translation; Theory: Predictions of judged motion direction by the Contrast-Strength-Vector-Sum theory with default parameters. The horizontal lines labeled C10, C20, represent the angles (relative to rigid direction) of the 10, 20Hz sinewave 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, 359 deg; 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 sinewaves. The plaid thumbnails at the top indicate (from left to right): the pure 20Hz stimulus, a plaid with equal-contrast 10 and 20 Hz components, and the pure 10Hz stimulus. The curve parameter (a) indicates component frequencies 10Hz+20Hz (200msec), and 4 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: (b) 10Hz+5Hz (650 msec), (c) 2Hz+1Hz (1,000 msec). Theory: The 11 different curves represent predictions as $\alpha$ (the proportion of third-order motion processing versus first-order) varies from 0 (pure first-order motion, top) to 1 (pure third-order motion, bottom).
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 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 sinewave components. Component contrast strength increases as 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 2 plaids towards the direction of rigid translation. For component temporal frequencies of 1 and 2 Hz with stimuli in which both plaid components were of high contrast and differed in contrast by a factor of 2 or less, the perceived motion direction was precisely the direction of rigid translation. There was no influence of first-order motion, i.e., pure third-order motion. For low temporal frequency plaids with unequal component contrasts, the judged direction of motion deviated from rigid towards the direction of the dominant plaid component in proportion to the power of the difference in component contrasts, an alternative direction of motion also computed by the third-order motion system. For same-spatial-frequency plaids, using only default parameters, with zero parameters estimated from the data, the Contrast-Strength-Vector-Sum theory describes both the first-order and third-order motion-system properties, predicting perceived motion direction over the full ranges of temporal frequencies, contrasts, and angles between the plaid components.
REFERENCES


Sperling, G. (1990). Type 1 and Type 2 experiments. [http://www.cogsci.uci.edu/whipl/Type_1_and_Type_2_Expts.pdf](http://www.cogsci.uci.edu/whipl/Type_1_and_Type_2_Expts.pdf)


A plaid is the superposition of two sinewave gratings that move independently in different directions, with different speeds and contrasts. Five themes: 1. The components of velocity, direction and speed, are computed separately in early stages of motion processing; the focus here is exclusively on direction. 2. Procedure demonstration: When most stimuli occur in a restricted range of directions, the perception of direction is distorted, therefore motion
stimuli must occur in random directions, 0-360deg. 3a. When only first-order system is
stimulated (temporal frequencies ≥10Hz 2 Ss, ≥15Hz 1 S), over the entire range of visible
contrasts, the contrast ratio of the two plaid components completely determines perceived
direction. (3b) Each component sine wave is represented as a vector: direction (perpendicular
to the sinewave stripes), and length (a factor ρ representing the relative effectiveness of
that temporal frequency times sinewave contrast to a power β ≈ 2). Perceived direction is
determined by contrast-strength vector summation, velocity is irrelevant. (3c) Once the β
and ρ for a subject have been determined for a particular set of plaids, all with the same
component angle, the same parameters predict 97% of the variance of new data with plaids
composed of a full range of possible angles. 4. For slow (1&2Hz equal-high-contrast plaids,
exclusively third-order motion is perceived – movement in the direction of rigid translation
(pattern direction, intersection of constraints). At intermediate contrast ratios and temporal
frequencies, a combination of 1st & 3rd order motion is perceived. (Second-order motion
is irrelevant for these plaids). 5. A purely theoretical, zero-estimated-parameter theory
that embodies the above principles, captures the essence of the full range of the data for
same-temporal-frequency plaids.
Protocol and parameter details for Experiment 1.

Low-contrast Type 1 symmetrical plaids were presented to subjects in two types of sessions: moving in unrestricted random directions (0, 359) deg and restricted directions (-4, +4) deg. Each plaid consisted of two sinewave components, with equal or unequal contrasts. Each component had a spatial frequency of 1.0 cpd and a temporal frequency of 10.6 Hz. Five contrast-ratios were used, with the higher contrast fixed at 2%. Each pair of unequal contrasts was used in two mirror-opposite plaids. The direction of rigid motion is arbitrarily designated as zero degrees. If we use C1 to indicate the plaid component moving counterclockwise from rigid direction, and C2 clockwise, then the two mirrored plaids can be represented by (C1, C2) and (C2, C1). Results from (C2, C1) are pooled with results from (C1, C2) after flipping the former around zero degrees.

Experiment 4.

The section contains the details to support the conclusions reported in the text.

Method

Stimuli. Type 2 plaids were used: 10 and 20 Hz components, both 1 cpd with an angle of 22.3 deg between them (as in Experiment 2). The component with the higher contrast had a contrast of 32%. The ratios of the lower contrast component to the higher contrast component were: 1, 1/2, 1/4, 1/8 and 0. This aspect of the design is similar to that of Experiments 1 and 2 (e.g., see Figs. 3, 11).

Procedure. There were two conditions: motion, and static (Fig. 11). The motion condition was identical to the same conditions in Experiment 2. The stimulus was displayed for 200 msec in a random orientation 0-359 deg, and the subject made a judgment of motion...
direction 0,...,359 deg, as before. The static condition was similar to the motion condition
except in two respects: Only one "static" snapshot of the plaid was shown for 200 msec.
In the static condition, the subjects’ task was to judge the orientation of the main axis of
the static plaid, i.e., of the gray lines separating the high-contrast alternating bars or of the
orientation of the bars themselves (Fig. 11). The orientation response range was 0,...,180
deg.

In order to compare results of orientation judgment with that of motion-direction judg-
ment, the orientation perpendicular to that of subjects’ judgment was used to display the
results of the orientation judgment.

Two subjects who had participated in Experiment 2 also participated in Experiment 4.

Results

Figure 12 shows the data (both motion judgments and orientation judgments) for the two
subjects. Both subjects show similar patterns of judged orientation. For subjects ROJ and
FR, respectively, the PSE for (32%, 32%) component contrasts are 60.1 and 58.2 deg, very
close to the second-order orientation of 58.3 deg. For all other contrast ratios, the lower
contrast grating is virtually ignored, and the judged orientation is that of the 32% contrast
grating. This is significantly different from the second-order orientation which is independent
of contrast ratio.

Subject ROJ’s motion data Fig. 12B are generally similar to the orientation data Fig.
12A. For contrast ratios of 4:1 and greater, the higher-contrast grating dominates: it de-
termines both the perceived orientation and the perceived motion direction. When the
components are of equal contrast, the PSE for perceived motion direction is 63.3 deg, which
diffs slightly but significantly from the perceived orientation of 60.1 deg. The big differ-
ence between perceived orientation and motion direction occurs for gratings of 2:1 contrast
ratio. The lower contrast component is ignored in orientation judgements but exerts a very
significant impact on motion direction.
Figure 14: A comparison of orientation judgments in snapshots of Type 2 plaids with motion direction judgments for the same plaids with components now moving at 20 and 10 Hz. (A,C) Two subjects’ data from orientation judgment task. The directions perpendicular to subjects’ original orientation judgments are shown, because they would correspond to subjects’ motion direction response if a strategy of using the orientation of plaid pattern as a clue to infer motion direction were used. The contrast of 20 Hz component was 32%, the contrast of the 10 Hz component was varied as indicated. Except for the equal contrast components (32%, 32%), orientation judgments were entirely or almost entirely determined by the higher-contrast component. (B,D) Motion direction judgments. (C) Subject ROJ’s motion judgments differ from orientation judgments primarily for component gratings with a 2:1 contrast ratio. (D) Subject FT’s motion direction judgments with 20 and 10 Hz components, as in Experiment 2, deviate strongly towards the rigid direction and no longer lie between the component vectors. The two dots on in right middle of Panel D represent FT’s data from Experiment 2 with components of 30 and 15 Hz at contrasts, from left-to-right, of (32%, 16%) and (32%, 32%).
Subject FT’s motion data, which lie almost entirely outside of the angle between the two components, are strongly deviated towards the rigid direction; the motion judgments are completely different from the orientation judgments. A similar data pattern was observed for this subject in Experiment 2 and interpreted as showing a residual influence of third-order motion even at these high temporal frequencies. In Experiment 2, with components of 30 and 15 Hz, this subject show the same data pattern as did the other subjects at 20 and 10 Hz. Two points of the Experiment 2 motion judgments are plotted in Fig. 12D. These points are quite similar to the equivalent points for the other subject. Again, the big difference between the very high temporal frequency motion judgments and the orientation judgments occurs when the plaid components have a 2:1 contrast ratio.

It was noted above that orientation judgments cannot be the main contributor to motion-direction judgments because orientation is independent of component temporal frequency. This is illustrated in Fig. 12C and 12D, where motion direction judgments change tremendously as temporal frequency changes even as orientation remains invariant. At the highest temporal frequencies, when only the first-order motion system is activated, the perceived motion and perceived orientation differ greatly as a function of the contrast of the plaid components. The independence of orientation and temporal frequency, and the difference between perceived orientation and perceived motion direction as a function of component contrast ratio suggest a minimal role for perceived orientation in judgments of plaid motion direction.