Contrast gain control in first- and second-order motion perception

Zhong-Lin Lu

Department of Psychology, University of Southern California, Los Angeles, California 90089

George Sperling

Human Information Processing Laboratory, Department of Cognitive Sciences and Institute for Mathematical Behavioral Sciences, University of California, Irvine, Irvine, California 92697

Received October 23, 1995; revised manuscript received July 9, 1996; accepted July 11, 1996

A novel pedestal-plus-test paradigm is used to determine the nonlinear gain-control properties of the first-order (luminance) and the second-order (texture-contrast) motion systems, that is, how these systems' responses to motion stimuli are reduced by pedestals and other masking stimuli. Motion-direction thresholds were measured for test stimuli consisting of drifting luminance and texture-contrast-modulation stimuli superimposed on pedestals of various amplitudes. (A pedestal is a static sine-wave grating of the same type and same spatial frequency as the moving test grating.) It was found that first-order motion-direction thresholds are unaffected by small pedestals, but at pedestal contrasts above 1–2% (5–10× pedestal threshold), motion thresholds increase proportionally to pedestal amplitude (a Weber law). For first-order stimuli, pedestal masking is specific to the spatial frequency of the test. On the other hand, motion-direction thresholds for texture-contrast stimuli are independent of pedestal amplitude (no gain control whatever) throughout the accessible pedestal amplitude range (from 0 to 40%). However, when *baseline carrier* contrast increases (with constant *pedestal modulation* amplitude), motion thresholds increase, showing that gain control in second-order motion is determined not by the modulator (as in first-order motion) but by the carrier. Note that baseline contrast of the carrier is inherently independent of spatial frequency of the modulator. The drastically different gain-control properties of the two motion systems and prior observations of motion masking and motion saturation are all encompassed in a functional theory. The stimulus inputs to both first- and second-order motion process are normalized by feedforward, shunting gain control. The different properties arise because the modulator is used to control the first-order gain and the carrier is used to control the second-order gain.

1. INTRODUCTION

The human visual system functions over an enormous range of input light levels extending from extremely dim starlight ($\sim 10^{-3}$ cd/m^2) to very bright sunlight ($\sim 10^5$ cd/m^2). However, most visual phenomena are independent of absolute luminance level for an extremely wide range of luminances.[1] This is largely accomplished by preceding all the other visual processes with a mechanism of retinal adaptation[2–8] that removes the mean luminance from the visual input and provides, to subsequent processes, only contrast—the fraction of the relative increase or decrease (with regard to the mean luminance of its neighborhood) of input light at each point in space.

Following light adaptation in the visual system, there is contrast gain control.[9] Neurons in the visual pathway, beginning at the level of retina ganglion cells[8] and continuing in the lateral geniculate nucleus[10,11] and the primary visual cortex,[12–18] all demonstrate some degree of contrast gain control. Their responses do not increase with the input contrast beyond a certain level. The functional significance of this is that, once a stimulus achieves a critical level of contrast further increases in contrast do not affect the neural representation. It allows the brain to compute certain kinds of stimulus information without the distraction of irrelevant contrast variations. For example, judgments of the distance between two objects or of an object's velocity clearly would benefit from being independent of object contrast because contrast is irrelevant once it is sufficient to make the objects clearly visible. At the perceptual level, numerous visual tasks have indeed been shown to be independent of contrast[19–24] for contrasts above approximately 5–10%. In this paper, we are concerned with the contrast gain-control properties of the human visual motion system.

Much of our knowledge about the human perceptual motion mechanisms has been derived from psychophysical experiments at extremely low contrast levels.[25–31] At very low contrasts (e.g., less than 2%) one assumes that processing stages before motion extraction behave linearly with respect to motion stimuli, and therefore experiments with such stimuli focus on the nonlinear properties of the motion and subsequent decision mechanisms. However, in normal daily life we are confronted with perceptual tasks that involve the full range of contrasts. To completely describe or simulate human behavior in a natural environment, one must find ways to extend our knowledge obtained in near-threshold conditions to higher-contrast conditions. Studies of spatial and temporal contrast masking in pattern vision[32–35] suggest that contrast gain control plays a major role in pattern vision.

Because it has been well documented that different visual pathways have distinct gain-control properties,[11,17,36] we might expect contrast gain control in motion perception to be different from contrast gain control in other systems; and beyond that, we might expect different gain-control properties for different motion perception mechanisms.

We know of two studies of contrast gain-control (saturation) properties of the first-order (luminance-modulation) motion system.[20,37] These are incomplete in various ways, and we consider them in some detail in Section 5. We know of no study of contrast gain control in the second-order (contrast-modulation) motion system. It is now believed that there are at least three quite different motion mechanisms.[31,38] Here we offer an improved paradigm that enables us to isolate the first-order and the second-order motion mechanisms and to study the contrast gain-control properties of each independently.

2. THEORY

A. Reichardt and Motion-Energy Detectors

The Reichardt detector (Fig. 1), a basic motion-extraction unit in computational theories of motion perception, was originally developed for insect vision by Reichardt[39,40] and was successfully adapted by van Santen and Sperling[30] for human perception. It consists of two mirror-image subunits [e.g., left (L) and right (R)] tuned to opposite directions of motion. Subunit L multiplies the signal at spatial location A with the delayed signal from a rightward adjacent spatial location B. Subunit R multiplies the signal at spatial location B with the delayed signal from spatial location A. The output of each subunit is integrated for a period of time, and the direction of movement is indicated by the sign of the difference between the subunit outputs.[39,40] In terms of the overall system input–output transformations, Reichardt detectors can be proved[41] to be equivalent to (a) motion energy detectors[42] and (b) motion filters based on Hilbert transforms[43] after they have been elaborated into a motion detector. Thus there is no loss of generality in considering Reichardt detectors.

B. First-Order and Second-Order Motion

When a Reichardt detector is applied to the raw or linearly filtered visual input, this application is called a first-order analysis.[44,45] Indeed, first-order analysis provides an accurate account of motion direction for an enormous range of stimuli,[30,43,46–48] including some quite complex waveforms and quite counterintuitive predictions.[30] However, Chubb and Sperling[49,50] demonstrated clear motion perception in broad classes of (drift-balanced and microbalanced) stimuli constructed of drifting modulations of contrast, spatial frequency, texture type, or flicker (see also Refs. 51–55) whose motion was completely invisible to Reichardt detectors. Such stimuli were said to activate second-order motion mechanisms[44,45] because Chubb and Sperling noted that spatiotemporal filtering plus some gross nonlinear preprocessing (e.g., absolute value or square-law rectification) before a Reichardt detector could expose the latent motion in drift-balanced and microbalanced stimuli.

C. Pedestal Immunity of Reichardt Detectors

Van Santen and Sperling[30] proved mathematically that, for continuous, infinite-duration motion stimuli, Reichardt (and the equivalent motion-energy) detectors have two remarkable properties: (1) Pseudolinearity: The detector's output to the sum of several sine waves with different temporal frequencies is the sum of its outputs to the individual sine components (this would not necessarily hold if the components had the same temporal frequencies—therefore only *pseudo*linearity). (2) Ignoring static sinewaves: The Reichardt detector's output for any static sine-wave input—indeed, any static input—is zero.

Consider a pedestaled motion stimulus (Figs. 2c and 3c), that is, a compound stimulus resulting from linear superposition of a drifting sine wave (the motion stimulus) and a static sine wave of the same spatial frequency (the pedestal). A corollary from the properties of pseudolinearity and the ignoring of static displays is that the output of an elaborated Reichardt detector to a pedestaled stimulus is exactly the same as its output to the motion component alone. This is the *pedestal immunity* of Reichardt detectors.

The theorem that Reichardt detectors have pedestal immunity with infinitely long-duration continuous stimuli can be extended to temporally sampled, finite-duration motion stimuli.[31] For a regularly sampled, finite-duration stimulus, pedestal immunity holds asymptotically when the following conditions are satisfied:

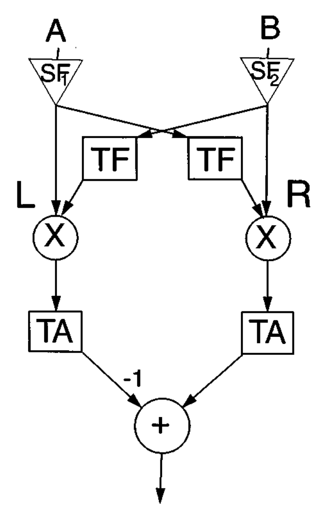

Fig. 1. Elaborated Reichardt detector. It computes motion direction from two inputs that sample the visual display at two adjacent spatial locations A and B. SF_1 and SF_2 denote linear spatiotemporal filters (receptive fields) that may be different from each other. In the right (R) subunit of the detector, the output of SF_1 is delayed by the temporal delay filter TF and then multiplied ($\times$) by the direct output of SF_2. The output of the multiplier is temporally averaged over a temporal window (defined by a linear filter TA) to produce the final output of the R subunit. In the L subunit of the detector, the output of SF_2 is delayed by TF, multiplied ($\times$) by the direct output of SF_1 and temporally averaged by TA. The difference (R minus L) defines the output of the detector. Outputs greater than zero indicate stimulus motion from A to B; outputs less than zero indicate stimulus motion from B to A.

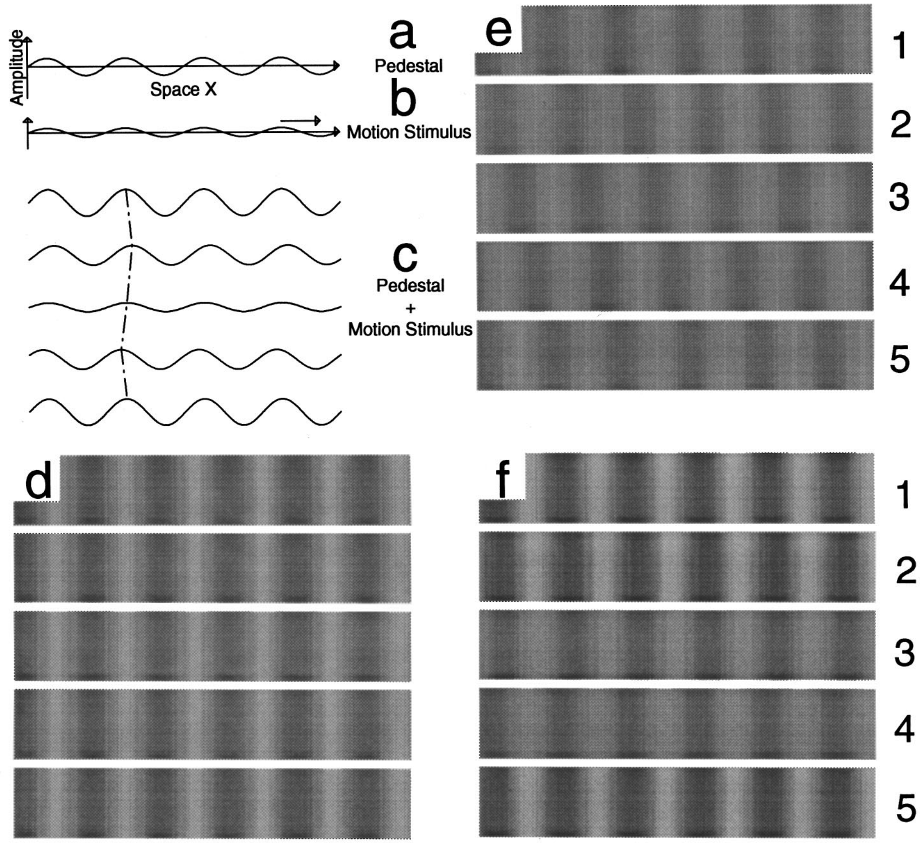

Fig. 2. First-order (luminance-modulation) motion stimuli, with and without pedestals. (a) The pedestal: a static sine wave. (b) The motion stimulus: a moving sine wave with half the amplitude of a, frozen at one instant in time. (c) Five frames of pedestaled motion: the sum of (a) and (b). From frame to frame the moving sine wave travels 90 deg to the right. The wobbling movement of a peak of the compound waveform is indicated by the dotted–dashed line. Any mechanism that computes motion from stimulus features such as peaks, valleys, or zero crossings perceives only the wobble. (d, e, f) The five stimulus frames shown separately for each component. Each horizontal segment (1–5) shows a slice of the component as it would have appeared to the subject, except that the contrast has been enormously exaggerated for the purpose of reproductive clarity. (d) The static sine-wave pedestal, as diagrammed in (a). (e) The drifting luminance modulation, as in (b); consecutive frames are shifted to the right by 90 deg. (f) Pedestaled motion: the sum of modulations (d) and (e). The five frames correspond to those shown schematically in (c).

(1) the stimulus lasts one full temporal cycle plus one extra frame; (2) the time constant of the output filter (TA in Fig. 1) in the Reichardt model is long relative to a stimulus cycle.

D. Motion Pedestal Test

The pedestal immunity of Reichardt detectors can be used to determine whether motion perception is compatible with a Reichardt algorithm.[30,31] If an observer computed motion direction by means of a Reichardt detector, the observers' performance would be the same when the motion stimulus is shown alone as when it is pedestaled.

The pedestal test is a powerful paradigm for distinguishing between different motion-extraction mechanisms, because pedestal immunity is a rather unusual property. For example, a pedestaled stimulus with a pedestal:test amplitude ratio of 2:1 (Figs. 2c, 2f, 3c, and 3f) is made of sine waves of the same spatial frequency with a back-and-forth phase oscillation (across frames) equal to 1/6 of the spatial cycle. If the motion direction computation were based on stimulus features (peaks, valleys, zero crossings,[56] etc.), the pedestaled stimulus would appear to wobble, and it would be impossible for subjects to judge the motion direction of the test component.

E. Pedestal Immunity of Human Observers: Small Pedestal Amplitudes

With small total modulation depth of the pedestaled stimulus, subjects' performance in motion-direction discrimination is the same for pedestaled and nonpedestaled motion stimuli.[30,31] This pedestal immunity holds exactly for first-order (luminance-modulation) and second-order (texture-contrast-modulation) motion.[31] And pedestal immunity holds exactly whether the duration of the pedestal and motion stimuli together is extremely brief (1/16 s) or of any longer duration, so that selective temporal filtering of the pedestal is impossible and therefore cannot account for pedestal immunity. Pedestal immunity and several equally counterintuitive properties all add support to the hypothesis that Reichardt/motion-energy detectors are used in these motion computations.[30]

F. Pedestal Immunity of Human Observers: Large Pedestal Amplitudes

With large pedestal amplitudes, the pedestal immunity of human observers can break down (see below). We con-

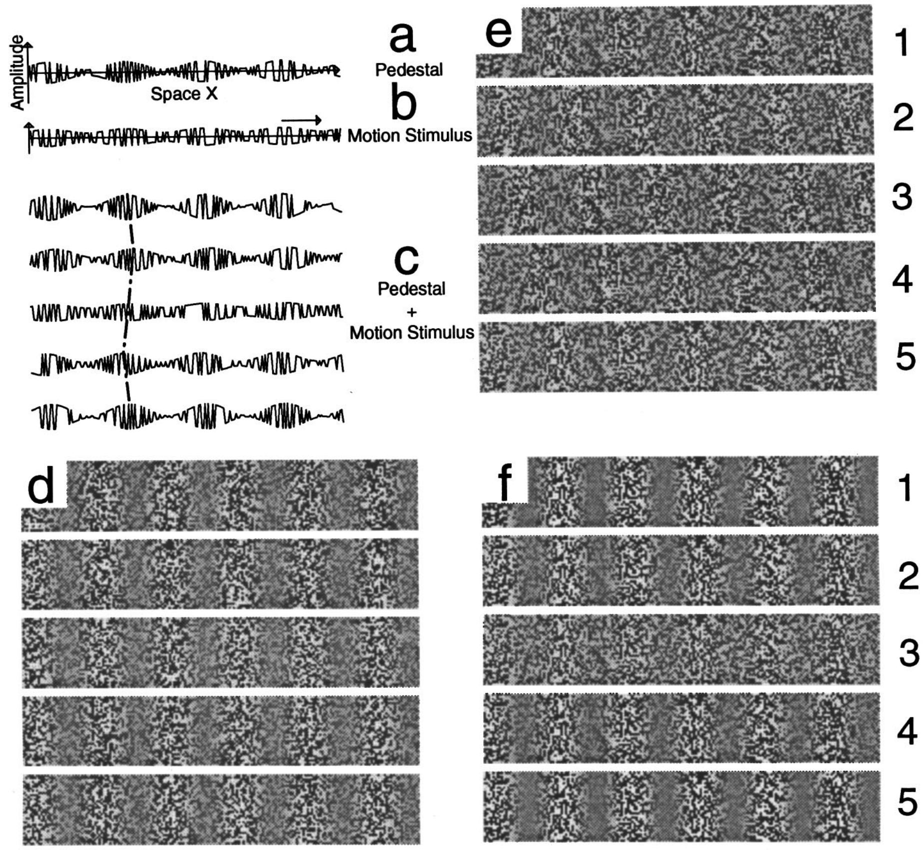

Fig. 3. Second-order (texture-contrast-modulation) motion stimuli, with and without pedestals. (a) The pedestal: a static sinusoidal modulation of texture contrast. (b) The motion stimulus: a moving sinusoidal texture-contrast modulation with half the amplitude of (a) frozen at one instant in time. (c) Five frames of pedestaled motion: the sum of (a) and (b). From frame to frame the moving sinusoidal modulation travels 90 deg to the right. The wobbling movement of a peak of the compound waveform is indicated by the dotted–dashed line. Any mechanism that computes motion from stimulus features such as peaks, valleys, or zero crossings perceives only the wobble. (d, e, f) The five stimulus frames shown separately for each component. Each horizontal segment (1–5) shows a slice of the component as it would have appeared to the subject, except that the contrast has been exaggerated for the purpose of reproductive clarity. (d) The static sine-wave pedestal, as diagrammed in (a). (e) The drifting texture-contrast modulation, as in (b); consecutive frames are shifted to the right by 90 deg. (f) Pedestaled motion: the sum of modulations (d) and (e). The five frames correspond to those shown schematically in (c).

ceptualize the breakdown of pedestal immunity as a nonlinear gain-control process that occurs before the motion computation itself. The alternative, that gain control occurs within the motion computation itself is not excluded, but it is computationally and conceptually much less attractive, and it is not necessary. In the study of amplitude saturation in the motion system, pedestal amplitude is the independent variable. We determine the threshold modulation of first-order (Fig. 2) and of second-order (Fig. 3) motion stimuli (each on its own type of pedestal) for pedestal amplitudes over a wide range.

3. GENERAL METHOD

A. Stimuli

Most visual phenomena in the study of motion perception are relatively independent of the absolute luminance level for an extremely wide range of luminances.[1] Therefore it is convenient to define stimuli $s(x, y, t)$ in terms of their point contrast:

$$s(x, y, t) = \frac{L(x, y, t) - L_0}{L_0}, \tag{1}$$

where $L(x, y, t)$ is the luminance at the point (x, y, t) and L_0 is the mean luminance of the display area.

All the contrast functions $s(x, y, t)$ considered here can be described as the product of a modulation function $M(x, t)$ and a static carrier $C(x, y)$ with the following properties:

1. The carrier is defined within a display window that is surrounded by a uniform background. The expected luminance is the same across the entire display.
2. For the luminance-modulation stimuli, $C(x, y) = 1$; for the texture-contrast-modulation stimuli, the expected mean contrast $E[C(x, y)] = 0$, but $C(x, y)$ is a random variable that takes $+1$ or -1 with equal probability.
3. The modulator consists of linear summation of (a) a baseline m_b, (b) a pedestal (a static sine wave with modulation depth m_p), and (c) a drifting sine-wave grating with modulation depth m. The static pedestal has spatial frequency α, temporal frequency 0, and phase θ_p; it is $m_p \sin(2\pi\alpha x + \theta_p)$; the drifting sine wave is defined by $m \sin[2\pi(\alpha x + \beta f t) + \theta]$. It moves leftward when

$\beta = 1$ and rightward when $\beta = -1$. It has modulation depth m, initial phase θ, spatial frequency α, and temporal frequency f. The complete modulator is given by

$$M(x, t) = m_{\mathrm{b}} + m_{\mathrm{p}} \sin(2\pi\alpha x + \theta_{\mathrm{p}}) + m \sin[2\pi(\alpha x + \beta f t) + \theta]. \quad (2)$$

4. The modulator was regularly sampled four times during a stimulus cycle so that the phase shift of the motion component between successive frames was $\pi/2$ (90 deg). The duration of the modulator was five frames (the initial frame was repeated as the final frame).

B. Apparatus

All the displays were created in advance of a session by HIPS—an image-processing software package[57,58] on a Unix computer (SunSparc II). Displays were presented on an IKEGAMI DM516 achromatic graphics monitor with a 20-in. diagonal screen, a P4 (fast white) phosphor, and 60-Hz vertical retrace, driven by a TrueVision AT-Vista video graphics adapter residing in an IBM-486PC-compatible computer. The C-language programs that controlled display scheduling and data collection were based on the Runtime Library software package.[59]

The graphics system produces 4096 (12 bits) distinct gray levels with a dynamic range of 12.1 cd/m^2 (when every pixel is assigned the lowest gray level) to 325 cd/m^2 (when every pixel is assigned the highest gray level). The background luminance was made equal to the midpoint $L_0 = 0.5(325 + 12.1) = 169$ cd/m^2. The point-contrast range of such displays is -0.93 to $+0.93$. A psychophysical calibration procedure was used to linearly divide the whole luminance range into 256 gray levels. When contrasts below 1% were required, a second lookup table was generated to create 16 gray levels between contrasts of ± 0.0073.

All the displays were viewed binocularly with natural pupil at a viewing distance of 114 cm in a dimly lighted room (the average luminance of surfaces in the room was approximately 10 cd/m^2). The stimuli occupied the central 6.34 deg $\times$ 3.17 deg of a uniformly luminous CRT screen (L_0, 169 cd/m^2) that extended 17.1 deg $\times$ 11.2 deg.

C. Procedure

The purpose of the procedure was to determine the threshold modulation depth m_{th} for correctly judging motion direction (L versus R) for each subject and each stimulus condition. The method of constant stimuli[60] was used to generate psychometric functions. The 75%-correct point was estimated from each psychometric function. Psychometric functions consisting of five points were obtained for the two types of motion stimulus (luminance and texture-contrast), for each of the pedestal amplitudes and for the no-pedestal condition for different texture baseline contrasts and for each subject. At least 80 observations were made by each subject at every point on the psychometric functions.

D. Trials

The subject initiated a trial by pushing a button. A fixation point appeared immediately at the center of the display and remained on throughout the trial. It was followed in 0.5 s by the motion stimulus and the pedestal, which started concurrently, each with a random one of four phases. The stimulus presentation consisted of five frames (a full temporal cycle plus one extra frame) of the motion stimulus. The extra frame was added so that the last frame was always identical to the first frame. This removed any positional cue on which subjects could base their judgments and served to meet the sampling condition under which pseudolinearity holds for Reichardt detectors.[31]

The subject's task was to judge direction of movement (L or R). The judgment was made by pushing one of two buttons. The percent of correct (as defined *a priori* by the experimenter) judgments of motion direction was the main dependent variable of all the experiments. Feedback was given to the subject immediately after each trial.

The three experiments all determined modulation-amplitude thresholds of the motion stimulus for 75%-correct motion-direction judgments. Experiment 1 measured luminance thresholds m_{th_1} as a function of the amplitude m_{p_1} of a luminance-modulation pedestal. Experiment 2 measured texture-contrast modulation threshold m_{th_2} as a function of the amplitude m_{p_2} of a texture-contrast-modulation pedestal. Experiment 3 measured the threshold amplitude m_{th_3} of a moving texture-contrast modulation as a function of the baseline texture-contrast modulation m_{b_3}. In each experiment, all stimulus conditions (pedestal amplitude m_{p} or baseline amplitude m_{b}, pedestal phase θ_{p}, motion stimulus modulation depth m, motion starting phase θ, and motion direction β) were mixed within experimental blocks. A block normally consisted of about 360 trials and lasted approximately 20 m. Intermissions between blocks were ~5 min. Subjects normally were given a 2-min dark-adaptation period if they entered the test room from daylight. A session lasted approximately 2 h.

E. Subjects

A University of California, Irvine, graduate student (EB), naïve to the purposes of the experiments, and the first author served as subjects in all the experiments. Both are male and have corrected-to-normal vision.

4. EXPERIMENTS

A. Experiment 1. Contrast Gain Control with First-Order Stimuli: Motion Threshold m_{th_1} as a Function of Pedestal Amplitude m_{p_1}

A first-order (luminance) stimulus was made of a rigidly drifting sine-wave grating (the motion stimulus) superimposed on a static sine-wave grating of the same spatial frequency (the pedestal). This is the kind of first-order stimulus whose motion direction can be computed directly by a Reichardt/motion-energy mechanism.[30,31] It had been conjectured by Nakayama and Silverman[20] that the luminance modulation motion system saturates at fairly

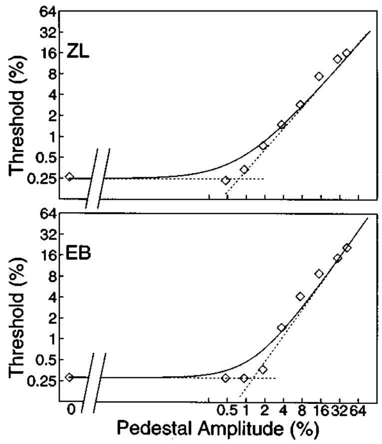

Fig. 4. Contrast-modulation threshold for 75%-correct motion-direction judgments versus pedestal-modulation amplitudes for first-order stimuli. Both axes are logarithmic. Each panel represents a different subject. The smooth curve drawn though the data represents Eq. (7f) in the text: The level of the horizontal asymptote is k', the threshold without a pedestal (0.25% and 0.28% for subjects ZL and EB, respectively). At large pedestal amplitudes, motion threshold for drifting luminance modulation is approximately proportional to the amplitude of the pedestals: η' is the slope of the diagonal asymptote (1.01 and 1.25 for subjects ZL and EB, respectively). The intersection of the asymptotes at $m_{\text{ped}} = (\lambda'/k')^{-\eta'}$ is the reciprocal of the generalized Weber fraction (the number of times that pedestal amplitude exceeds threshold amplitude: 0.75% and 1.2% for subjects ZL and EB, respectively).

low contrasts (approximately 4–5%). However, in reexamining their theory, we found serious problems, which we consider in Section 5. Here we offer a different approach (the motion-plus-pedestal paradigm) to the study of motion gain control in the first-order (luminance) motion system.

In Experiment 1 we measure motion threshold m_{th_1} as a function of pedestal amplitude m_{p_1} for luminance-modulation stimuli. The notion is that the static pedestal activates only the contrast gain-control mechanism, which then reduces the effective input to the motion system.

1. Method

Luminance-modulation stimuli are completely described by the modulator function:

$$M_1(x, t) = m_{\text{b}_1} + m_{\text{p}_1} \sin(2\pi\alpha x + \theta_{\text{p}_1}) + m_1 \sin[2\pi(\alpha x + \beta f t) + \theta], \quad (3)$$

where the baseline amplitude m_{b_1} is 0, the pedestal amplitude m_{p_1} is 0, 0.0047, 0.093, 0.019, 0.037, 0.074, 0.15, 0.30, or 0.42, the pedestal phase $\theta_{p_1} = 0$, $\pi/4$, $\pi/2$, or $3\pi/4$ with equal probability, and the initial phase of the motion stimulus θ is 0, $\pi/2$, π, or $3\pi/4$ with equal probability. The spatial frequency α of the sinewaves is 1.26 cycles per degree (c/deg) and the temporal frequency f is 7.5 Hz. Successive frames of the motion component are separated by $\pi/2$. A stimulus presentation consists of five frames (one full cycle plus one frame). The phase shift of $\pi/2$ between adjacent frames vitiates any possible second-order contribution to the motion computation. (The full-wave rectification of second-order motion would double the spatial frequency; thereupon the phase shift of $\pi/2$ would become π, and that would produce only counterphase flicker, not motion.) The temporal frequency of the stimulus was chosen to be twice the cutoff frequency of the third-order (feature-salience) motion system, so that third-order contributions also would be insignificant.[31]

2. Results

Figure 4 shows a log–log plot of the 75% threshold modulation m_{th_1} for discriminating motion direction as a function of pedestal amplitude m_{p_1}. For both subjects, motion threshold is constant for pedestal amplitudes of less than ~2%, and increases in direct proportion to pedestal amplitude for larger pedestals. We now consider these results in more detail.

For subject ZL, $m_{\text{th}_1} = 0.0025$ for $m_{\text{p}_1} < \approx 0.01$, and it increases with a slope of 1.012 for larger pedestals (a slope that is not significantly different than 1). For subject EB, $m_{\text{th}_1} = 0.0028$ when $m_{\text{p}_1} < \approx 0.015$, and it increases with slope of 1.25 for larger pedestals. A pedestal amplitude of 0.015 is clearly visible and is ~7× above its own threshold. This pedestal amplitude is more than 5× the threshold amplitude of the moving sine, so that the back-and-forth wobble of the pedestaled stimulus is less than 1/16 of a spatial period. Such a pedestal is quite sufficient to camouflage the linear movement. The first impression of the pedestaled stimulus is of a stationary pedestal; it requires a little practice to determine the direction of the moving component.

For small pedestals, these results complement our previous finding[30,31] that subjects' judgments of motion direction are equally accurate whether the motion stimulus is shown alone or is pedestaled. This finding argues strongly that motion direction of drifting luminance modulation is computed by the Reichardt class of detectors. We infer that when the total modulation depth is less than ~0.015, first-order motion signals are transmitted linearly before a Reichardt motion computation.

For pedestal amplitudes greater than ~3.7%, motion threshold is proportional to the amplitude of the pedestal, a Weber law. For subject ZL, $m_{\text{th}_1} = 0.39 m_{\text{p}_1}$ ($r^2 = 0.9930$, $n = 5$). For subject EB, $m_{\text{th}_1} = 0.49 m_{\text{p}_1}$ ($r^2 = 0.9973$, $n = 5$). Such large Weber constants (0.39, 0.49) indicate that the first-order motion system suffers severe nonlinear gain control. This nonlinear gain-control process is spatial-frequency specific: Lu and Sperling[31] found that adding large-amplitude stationary white noise had no effect on subjects' motion-direction-judgment performance. We defer further discussion of these results to Section 5.

B. Experiment 2. Contrast Gain Control with Second-Order Stimuli: I. Motion Threshold m_{th_2} as a Function of the Amplitude m_{p_2} of a Texture-Contrast Pedestal

Drifting texture-contrast modulation is a kind of second-order stimulus whose motion is not directly accessible to Reichardt motion-energy computations. The most plausible way to expose the latent motion signal in the stimu-

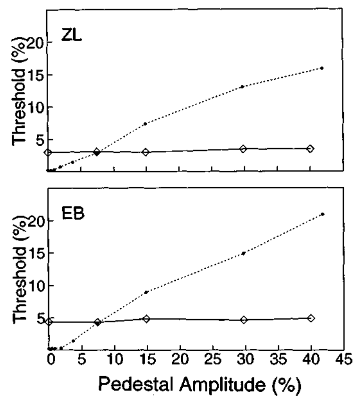

Fig. 5. Contrast-modulation threshold for 75%-correct motion-direction judgments versus pedestal-modulation amplitudes for second-order stimuli. Both axes are linear. Throughout the investigated pedestal-amplitude range, motion threshold for a drifting texture-contrast grating is constant for both subjects (diamonds connected by solid lines). Data from Experiment 1 are shown for comparison (dots connected by dotted lines).

lus is to preprocess the input with linear spatiotemporal filtering followed by rectification (e.g., absolute value or square-law) before a Reichardt-class detector.[45] Here we are concerned with the question of whether, in addition to rectification, there also is a compressive, saturating nonlinearity before motion detection. To answer this question, we measured motion threshold m_{th_2} as a function of pedestal amplitude m_{p_2} for texture-contrast-modulation stimuli.

1. Method

To make a drifting texture-contrast stimulus, the contrast of the carrier (a stationary, random, binary noise texture) is multiplied by the modulator (one plus a drifting sine wave). A pedestaled moving texture-contrast stimulus is created by linearly adding a static sine wave to the moving modulator before multiplying the modulator and carrier. The modulation function is

$$M_2(x, t) = m_{\mathrm{b}_2} + m_{\mathrm{p}_2} \sin(2\pi\alpha x + \theta_{\mathrm{p}_2}) + m_2 \sin[2\pi(\alpha x + \beta f t) + \theta], \qquad (4)$$

where the baseline amplitude m_{b_2} is 0.47; the pedestal amplitude m_{p_2} is 0, 0.074, 0.15, 0.30, or 0.41; the pedestal phase θ_{p_2} is 0, $\pi/4$, $\pi/2$, or $3\pi/4$ with equal probability; and the initial phase of the motion stimulus θ is 0, $\pi/2$, π, or $3\pi/4$ with equal probability. The spatial frequency of the modulator α is 1.26 c/deg, and its temporal frequency f is 7.5 Hz. The motion component moves $\pi/2$ between frames. A stimulus presentation consists of five frames (one complete period plus one additional frame).

The texture-contrast modulation is a pure second-order stimulus: Its expected luminance is the same everywhere; its motion cannot be determined by motion-energy detection,[45] and the fundamental Fourier motion components are useless.[50]

2. Results

Figure 5 depicts the 75% threshold modulation m_{th_2} for discriminating motion direction as a function of pedestal amplitude m_{p_2} for both subjects. The maximum physically possible range of pedestal amplitudes (with constant baseline contrast) is 0.47; the maximum pedestal amplitude range obtainable on our apparatus was 0.40 (we need the extra 0.07 to add motion stimuli). Throughout the investigated pedestal amplitude range ($0.0 < m_{\mathrm{p}_2} < 0.40$), motion threshold was virtually constant for both subjects ($m_{\mathrm{th}_2} = 0.032$ for subject ZL, and $m_{\mathrm{th}_2} = 0.047$ for subject EB). (For comparison, the very large changes in first-order thresholds are shown on the same graphs.) It is apparent that the texture-contrast motion system is immune to pedestals throughout the entire, very large, physically attainable amplitude range.

Immunity to pedestals complements our previous finding[31] that subjects' performance in judging motion direction of drifting texture-contrast modulation is the same whether the motion stimulus is shown alone or pedestaled. In second-order motion there is no restriction to small pedestals: There is pedestal immunity over the entire range of pedestal amplitudes. This argues strongly that motion direction of drifting texture-contrast modulation is computed by Reichardt-class detectors acting on the texture stimulus.

C. Experiment 3. Contrast Gain Control with Second-Order Stimuli: II. Motion Threshold m_{th_3} as a Function of the Baseline Contrast Amplitude m_{b_3} of a Texture-Contrast Pedestal

Experiment 2 established that threshold amplitude for judging motion direction of drifting texture-contrast modulation does not change with increasing pedestal amplitude. Does this mean that there is no contrast gain control in the texture-contrast motion system?

The texture-contrast pedestal, as described by Eq. (4), consists of a static baseline m_{b_2}, a static pedestal $m_{\mathrm{p}_2} \sin(2\pi\alpha x + \theta_{\mathrm{p}_2})$, and a drifting sine-wave grating $m_2 \sin[2\pi(\alpha x + \beta f t) + \theta]$. In Experiment 2, only pedestal amplitudes m_{p_2} were varied; the baseline amplitude m_{b_2} was kept constant. This made the expected contrast energy (over the whole stimulus) a constant, independent

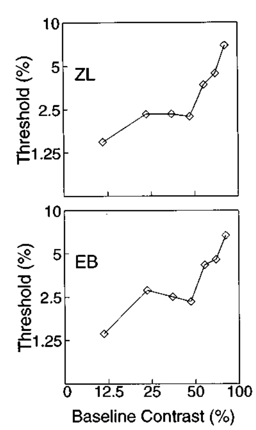

Fig. 6. Threshold for 75%-correct motion-direction judgments versus baseline contrast for second-order stimuli. There is no pedestal—no contrast modulation except for the motion stimulus. The coordinates are logarithmic. The horizontal axis denotes baseline root-mean-square amplitude.

of pedestal amplitude. The results of Experiment 2 were consistent with a contrast gain-control mechanism that takes as its control signal the total contrast energy from all the spatial-frequency channels combined. To explore this possibility in Experiment 3 we varied the baseline contrast m_{b_3} to determine its effect on motion-direction thresholds.

1. Method

The modulation function is described by Eq. (5):

$$M_3(x, t) = m_{\mathrm{b}_3} + m_{\mathrm{p}_3} \sin(2\pi\alpha x + \theta_{\mathrm{p}_3}) + m_3 \sin[2\pi(\alpha x + \beta f t) + \theta]. \quad (5)$$

All the stimulus parameters are the same as in Experiment 2, except that now pedestal amplitude m_{p_3} is zero and baseline amplitudes vary: $m_{\mathrm{b}_3} = 0.12$, 0.23, 0.35, 0.47, 0.58, 0.70, or 0.81.

2. Results

Figure 6 shows 75% thresholds for motion-direction judgments of texture-contrast stimuli as a function of baseline contrast m_{b_3}. For both subjects, thresholds seem to rise proportionally to baseline contrast (slope of ~ 1) except for a pronounced dip in the neighborhood of 47% baseline contrast. An intriguing conjecture is that the many thousands of trials the observers viewed at $m_{\mathrm{b}} = 47\%$ baseline contrast lowered their thresholds at 50% baseline level relative to their less-well-practiced performances at other contrasts. This issue cannot be resolved now. For the moment, we merely observe that gain control in second-order motion perception is closely related to the energy in the carrier texture and apparently has no dependence whatever on energy in the masking modulator, even though a test modulator of the same spatial frequency as the modulator is being detected.

5. GAIN-CONTROL MODEL OF MOTION SATURATION

The overall plan of this section is to show that a simple formulation[61] based on a feedforward shunting model of gain control, the sort originally proposed by Sperling and Sondhi[4] with the added feature that the control signals are first rectified,[9,18] accounts for the main results of these and similar experiments. Shunting inhibition is a concept that arises naturally from the mechanisms of neuronal inhibition,[62] and it results in divisive gain control (as opposed to subtractive inhibition). The neuronal mechanism of shunting inhibition is a decrease in the time constant of an RC circuit as a result of a decrease in R. Therefore it involves both a change in gain for low frequencies (gain = RC), and an increase in the cutoff frequency $(\mathrm{RC})^{-1}$, i.e., a greater proportion of high frequencies in the response. As only the gain change is of concern here, shunting inhibition can be represented by an amplifier that has an input $u(x, t)$ that is a function of space and time, as well as an output $v(x, t)$ that also is a function of space and time, and the amplifier has a variable gain that is controlled by a third function $w(x, t)$. For a gain-control amplifier at location x, the equation for shunting inhibition is

$$v(x, t) = \frac{u(x, t)}{k + w(x, t)}, \quad (6a)$$

where the positive real constant k is a threshold above which the gain control becomes effective. In the equations for shunting inhibition, k simply represents the resting conductance of the cell membrane when there is no input.[4]

To apply the basic shunting mechanism of Eq. (6a) to data, we need to specify the preprocessing of the visual stimulus that results in functions $u(x, t)$ and $w(x, t)$. The consecutive stages of processing are diagramed in Fig. 7a. These processing stages are described mathematically in the following paragraphs. Finally, there is a formal derivation that relates the processed signals to the statistics that are measured in the experiments. Thus the three elements that need to be brought into alignment are a block diagram representation, a mathematical representation, and a derivation of data statistics. Equation (6b) gives an overview of the preprocessing of the denominator of Eq. (6a) (the control signal), and Eq. (6c) gives an overview of the preprocessing of the numerator of Eq. (6a) (the test stimulus).

$$w(x, t) = \mathrm{Fc}[\mathrm{test}(x, t) + \mathrm{pedestal}(x, t)] \overset{1.}{\rightarrow} \mathrm{Fc}[\mathrm{pedestal}(x, t)], \quad (6b)$$

$$u(x, t) = \mathrm{test}(x, t) + \mathrm{pedestal}(x, t) \overset{2.}{\rightarrow} \mathrm{test}(x, t). \quad (6c)$$

In Eqs. (6), 1. and 2. indicate conditions that are described below. In Eq. (6b), Fc[ ] is a functional that involves (i) spatiotemperal filtering (convolution with a spatiotemporal impulse response, e.g., a receptive field), (ii) rectification (raising the absolute value of the signal at each point in space and time to a power η, $\|\cdot\|^{\eta}$), and (iii) integration over a space–time window. Because rectification is a nonlinear process, the computation implied by Fc[ ] can be quite complicated. However, the judicious choice of test stimuli and pedestals permits two major simplifications of Eqs. (6b) and (6c).

1. Both test(x, t) and pedestal(x, t) are spatial sinewaves. The test is a temporal sine wave, whereas the pedestal is simply turned on and then off (an all-positive temporal step function). Therefore the rectification operation in Fc[ ] doubles the temporal frequency of the test but not of the pedestal. These double and higher temporal frequencies average out in each half-cycle of the test, and therefore test(x, t), on the average, contributes zero to $w(x, t)$. So, although both test(x, t) and pedestal(x, t) appear within Fc[ ], we can ignore the test component and regard $w(x, t)$ as a function only of the pedestal(x, t).
2. Because of pedestal immunity, the response of the motion system depends only on test(x, t). Thus, in computing the effective input to the motion system, we ignore pedestal(x, t) in the numerator of Eq. (6a).

With the simplifications resulting from conditions 1 and 2 above, after contrast gain control, the effective input $v_{\mathrm{eff}}(x, t)$ to the motion system is simply

$$v_{\mathrm{eff}}(x,\ t) = \frac{\mathrm{test}(x,\ t)}{k + \mathrm{Fc}[\mathrm{pedestal}(x,\ t)]}. \tag{6d}$$

To relate Eqs. (6) to the data of Experiment 1, consider the statistics that were obtained from the experiment. The data are the estimated 75% threshold modulation amplitudes $m_{\mathrm{test},75}(m_{\mathrm{ped}})$ for sine-wave test$(x,\ t)$ stimuli as a function of the modulation amplitude of the pedestals m_{ped} on which they were superimposed. The sine-wave test is

$$\mathrm{test}(x,\ t) = m_{\mathrm{test},75} \sin(\alpha x \pm ft), \tag{7a}$$

a grating that drifts either leftward or rightward according to whether the sign in front of t is plus or minus, respectively. The pedestal is

$$\mathrm{pedestal}(x,\ t) = m_{\mathrm{ped}} \sin(\alpha x). \tag{7b}$$

The test and the pedestal of Eqs. (7a) and (7b), respectively, are defined only within a spatiotemporal window and are zero elsewhere.

Substituting Eqs. (7a) and (7b) into Eq. (6d) and noting that, at the 75% threshold in Experiment 1, $v_{\mathrm{eff}}(x, t) = v_{75} \sin(\alpha x \pm ft)$, we obtain

$$\begin{aligned} v_{\mathrm{eff}}(x,\ t) &= \frac{m_{\mathrm{test},75}(m_{\mathrm{ped}})}{k + \mathrm{Fc}[m_{\mathrm{ped}} \sin(\alpha x)]} \sin(\alpha x \pm ft) \\ &= v_{75} \sin(\alpha x \pm ft), \end{aligned} \tag{7c}$$

which simplifies to

$$v_{75} = \frac{m_{\mathrm{test},75}(m_{\mathrm{ped}})}{k + \mathrm{Fc}[m_{\mathrm{ped}} \sin(\alpha x)]}. \tag{7d}$$

The denominator can be simplified further: $\mathrm{Fc}[m_{\mathrm{ped}} \sin(\alpha x)] \approx \lambda m_{\mathrm{ped}}^{\eta}$. This is because the rectification operation in Fc[ ] (Figs. 7a and 7b) is a pointwise

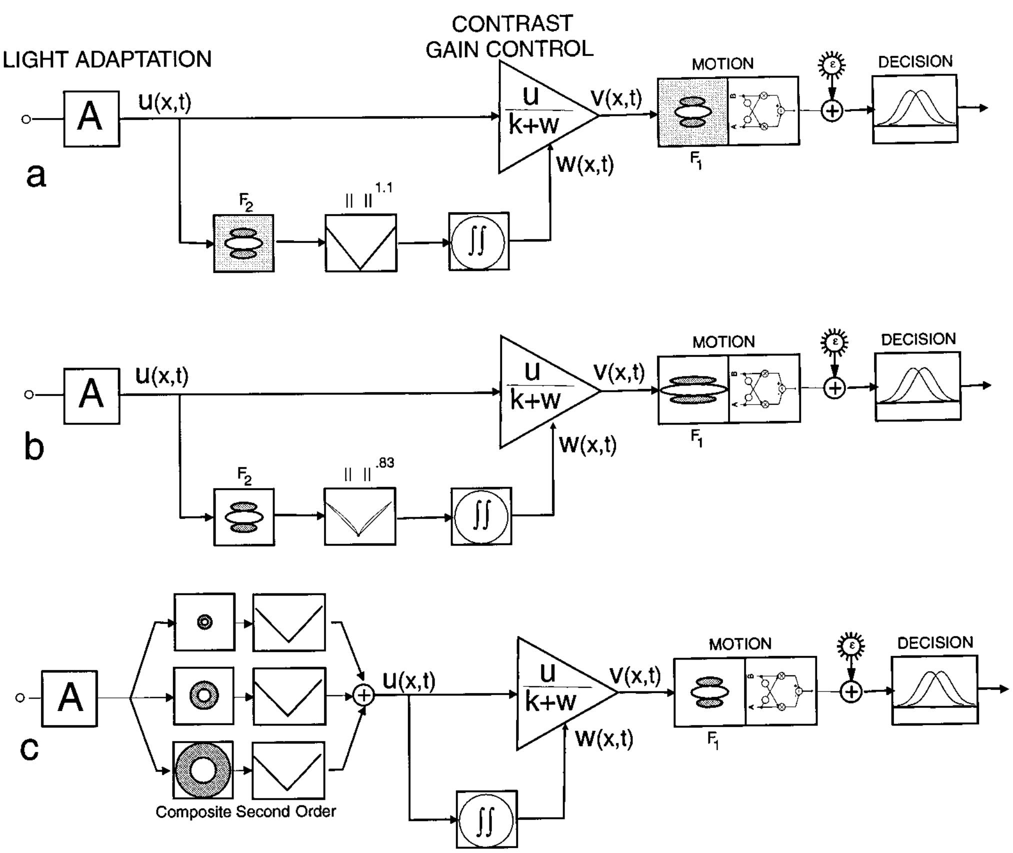

Fig. 7. Three models of motion gain control. In all the models the input signal first passes through a stage of light adaptation A and then divides into two paths which both arrive at a gain-control mechanism before the signal continues forward to the motion-detection component. (a) Model for first-order motion with pedestals derived from Experiment 1. The direct signal u arrives at the gain control, where it is divided by the controlling signal, $k+w$, which is first derived by filtering the input with a linear spatial filter F_2, rectifying the output (the actual exponent, the average of two subjects, $\eta = 1.1$), and then by forming a weighted sum to represent the spatial neighborhood of the motion detector ($\int\int$). The controlling signal $k + w(x,\ t)$ is the denominator of Eq. (6a) in the text. A Reichardt (motion-energy) detector receives the gain-controlled input $u(k + w)^{-1}$, and its output (support for rightward minus leftward motion) is subjected to additive noise ϵ and submitted to a decision mechanism. The linear filters in the gain-controlling F_2 and motion F_1 pathways are shown in gray to represent the fact that they are not constrained by Experiment 1. (b) Model representing the masking data of Anderson *et al.*[37] This model is identical to a except in the following details. The linear filter F_2 in the gain-controlling pathway is now defined by jittering mask stimuli, the linear filter F_1 in the motion pathway is now defined by moving mask stimuli, and the exponent of the rectification process is $\eta = 0.83$ (versus 1.1 for model a). (c) Model representing the second-order motion system as derived from Experiments 2 and 3. Radially symmetric spatial filters represent the center-surround receptive fields that define the range of spatial-frequency channels. Each filter is followed by rectification; the rectified channel outputs are summed (+) to produce the internal representation of the carrier plus modulator. The remainder of the model is as in (a) and (b)

nonlinearity. It takes the ηth power of the absolute value of its input. This rectification process doubles the spatial frequencies of sine-wave gratings $m_{\text{ped}} \sin(\alpha x)$. Subsequently, large receptive fields sum over several cycles, thereby minimizing variations of Fc[pedestal(x, t)] with x; so we drop the x. However, spatial summing also introduces an apparent η', which is different from η when $\eta \neq 1$. Spatial averaging over the sine pedestal after it has been raised to the η power is equivalent to simply taking the η' power of the peak amplitude where $\eta' < \eta$ when $\eta > 1$ and $\eta' > \eta$ when $\eta < 1$. Eliminating x and the sine components now yields

$$v_{75} = \frac{m_{\text{test},75}(m_{\text{ped}})}{k + \lambda m_{\text{ped}}^{\eta'}}. \tag{7e}$$

Because at 75% threshold v_{75} is a constant, it can be absorbed into two new constants $k' = v_{75}k$ and $\lambda' = v_{75}\lambda$, yielding

$$m_{\text{test},75}(m_{\text{ped}}) = k' + \lambda' m_{\text{ped}}^{\eta'}. \tag{7f}$$

In a log–log graph of $m_{\text{test},75}(m_{\text{ped}})$ versus m_{ped} (as in Fig. 4), Eq. (7f) describes the intersection of two asymptotes. The constant k' is the level of the horizontal asymptote, η' is the slope of the diagonal asymptote, and the asymptotes intersect at $m_{\text{ped}} = (\lambda'/k')^{-\eta'}$, the reciprocal of the generalized Weber fraction. The generalization of adding a slope parameter η' to Eq. (7f) has been sufficient to encompass data from a variety of quite different experiments.[8,9,18,34,47]

The plan of the remainder of this section is as follows: (1) to show that Eq. (7f) (generalized shunting gain control) accounts nicely for the results of Experiment 1 (first-order motion masked by pedestals); (2) to increase the scope of the model by considering pedestal and related masking paradigms in spatial vision with static stimuli; and (3) to consider motion stimuli that are masked by a variety of stimuli similar to but different from our pedestals. This enables us, in principle, to specify the model. (4) The first-order model is then shown to give a better, physiologically plausible account of two-flash experiments for which a quite different, nonphysiological model[20] had been proposed. (5) We then propose a model for second-order motion experiments that is analogous to the first-order model except that the gain is controlled by the carrier instead of by the modulator.

A. Minimal Model for First-Order Motion Pedestal Results

In all the models of Fig. 7, the input to the motion detector passes first through a stage of light adaptation (A), which divides the stimulus by the space–time average luminance value.[9] From this point on, the stimulus is represented by its local contrast, the normalized deviation of luminance at a point x, y from the average luminance. Subsequently, before any motion computation, the stimulus passes through a feedforward gain-control mechanism that is the present focus.

Figures 7a and 7c illustrate the models for gain control in the first- and the second-order motion systems, respectively, that were derived from the present experiments. Figure 7b illustrates a model for gain control derived by us from the motion-masking experiments of Anderson *et al.*,[37] which used quite different procedures. The first-order models in Figs. 7a and 7b are extremely similar. Both illustrate that quite different paradigms produce data that are accounted for by quite similar models and that each paradigm enables estimation of different components. The model derived from pedestal experiments is considered first.

1. The Model for Experiment 1

The model of Fig. 7a is derived from Experiment 1, in which static pedestals masked first-order motion stimuli. The spatiotemporal characteristics of effective masking stimuli are determined by a linear spatiotemporal filter, F_2, illustrated as a Gabor spatial filter. Such filters are approximations of the receptive fields of simple cells in area V1 of the occipital cortex. Simple cells occur with plus-center and minus-center receptive fields (i.e., positive and negative copies of filter F_2 as well as other phases), and they are distributed throughout space. In the brain, plus-center and minus-center neurons, like most neurons that communicate by means of neural firing rate, act like half-wave rectifiers, responding primarily positively by increasing their firing rate but being unable to reduce it below zero. In the brain, the outputs of plus-center and minus-center simple cells can be added to produce full-wave rectification or subtracted to produce an apparently linear response.[9] In the model of Fig. 7a, rectification is represented by the absolute value of the input raised to the power 1.1.

The rectified outputs of filters in the neighborhood of the motion detector are combined by means of a weighting function over space and time, indicated by the double integral. It is critical that the signal be rectified before integration. A moving sinusoid has a different phase at each different spatial location, which induces differently phased outputs of the filters at different spatial locations. If these differently phased filter outputs were simply to be added linearly over space and time, the conglomeration of positive and negative quantities would approximately cancel, and the expected value of such a sum would be zero. To produce a nonzero signal that indicates the total amount of stimulation, rectification must precede linear addition.

The motion detector (a Reichardt or equivalent motion-energy algorithm) computes the motion direction of its gain-controlled input. For sine-wave inputs, its output is proportional to m^2, the square of the sine amplitude.

In addition to noise that is inherent in the stimulus itself, the uncertainty of perceptual judgments is represented by the addition of noise ϵ to the signal before the decision process. The decision mechanism uses a simple real-valued criterion to decide whether the support for motion rightward exceeds the support for motion leftward.

2. Relating the Model to Data

The data from Experiment 1 are the mean threshold modulation amplitudes $m_{\text{test},75}$ of a drifting sinusoid as a function of the various pedestal amplitudes m_{ped} on which it is superimposed.

The constant η reflects a fundamental characteristic of the rectifier. When $\eta = 1$ then also $\eta' = 1$. However, for sine-wave stimuli, when $\eta < 1$ then $\eta' > \eta$, and when $\eta > 1$ then $\eta' < \eta$. This is not a large effect, but it does mean that the value of η in the model of Fig. 7b (0.83) is larger than the value of η' (0.73) in the data that produced it.

For the two observers in Experiment 1, the threshold value k' is 0.0025 and 0.0028. The constant λ' determines the knee of the threshold versus pedestal-intensity function in Fig. 4. In Experiment 1, λ' takes the values 0.35 and 0.70 for the two observers. For the data of Experiment 1, the slopes η' for the two subjects (Fig. 4) were 1.01 and 1.25, neither of which differed significantly from 1.0. Obviously, the rectification function is approximately linear, but the precise shape varies slightly from observer to observer. The average slope $\eta \approx \eta' \approx 1.1$ is indicated in Fig. 7a.

The model of Fig. 7a accounts for the main findings of Experiment 1. However, Experiment 1 contains no spatial variation in the configuration of masking or test stimuli, so it tells us nothing about the filters F_1 and F_2 nor about the parameters of the summation over space and time of the motion signal. Data that might determine the putative first-order motion components are available from several experiments[30,31,63,64] that measured absolute contrast thresholds for drifting sine waves. Very low contrast sine gratings are visible only to the first-order mechanism. However, because there undoubtedly are many channels of different spatial frequencies acting together, simple threshold data fail to define F_1; they merely define the properties of the mixture of filters that is active in a particular experiment. To learn more, we need to enlarge the scope of inquiry.

B. Pedestals Under Static Stimuli

We begin by considering the possible relevance to gain control in motion from what has been learned from the masking of static stimuli. For example, the pedestal method had been used in pattern recognition to study sine-wave detection in the presence of various masking sine waves.[32–34] For example, Legge and Foley[34] measured luminance contrast threshold of a 2-c/deg sine-wave grating in the presence of masking sine-wave gratings at various spatial frequencies and contrasts. One of their conditions, in which the masking and the test sine-wave gratings had the same spatial frequencies, is similar to our pedestal condition. They found that (1) at low contrasts (between 0.05% and 0.8%), a pedestal increases the detectability of the signal (a previously observed facilitation effect[32,33,65,66]) and (2) at higher contrasts (between 3.2% and 51.2%), the threshold $C_t = kC^{0.558}$, where k is a positive constant and C is the contrast of the masking sine-wave grating. It is apparent that the gain-control properties of pattern vision are somewhat different from those of the luminance motion system measured in Experiment 1. This seems to reflect a significant difference between form and motion vision.

C. Jittering and Moving Masking of First-Order Motion Stimuli

Anderson *et al.*[37] measured thresholds for judging motion direction of luminance-modulation sine waves under masks that varied in both spatial frequency and orientation. They measured motion threshold for a 3-c/deg drifting sine-wave luminance modulation in the presence of jittering sine-wave gratings of the same spatial frequency and various modulation depths m_p ranging from 0.075 to 0.42. They found that, in the masking contrast range under investigation, motion threshold was proportional to $m_{\mathrm{ped}}^{0.73}$ [Eq. (7f)]. When the jittering mask was replaced with a moving mask (motion-masking motion), there was increased masking and narrower spatial tuning. The model that we propose to account for their results (Fig. 7b) is identical in terms of the types and connections of components to the model proposed for our Experiment 1 (Fig. 7a). However, their extensive data with jittering pedestals can specify the filter F_2, and the comprehensive measurements of the additional masking produced by their motion maskers can specify filter F_1.

Anderson *et al.*[37] found that the masking by jittering masks increased with the 0.73 power of the input amplitude ($\eta' = 0.73$). We noted above that the exponent $\eta' = 0.73$ in the data implied that $\eta = 0.83$ in the process model. There are too many differences in stimulus configuration between the static pedestals of our Experiment 1 and the dynamically jittering maskers of Anderson *et al.* for us to know what might be responsible for the difference between exponents (0.83 versus 1.1). In this context, it would have been especially useful if Anderson *et al.* had studied mask contrasts below 0.075.

Anderson *et al.* found that moving masks have greater spatial-frequency selectivity than jittering masks. This tells us that the spatiotemporal selectivity of the motion detector (filter F_1 in Fig. 7b) is greater than that of the gain control mechanism (filter F_2 in Fig. 7b), something that could not be revealed by the procedure of Experiment 1. Anderson *et al.* also observed that moving masks are more effective masks than jittering masks. In the model of Fig. 7b, this is accounted for within the motion detector and the decision apparatus. Moving maskers exert a double masking effect: through the gain-control path and within the motion component itself. Extracting the signal motion from masked motion display ultimately is a problem that devolves on the decision/velocity mechanism. In the case of motion masking, the decision component must determine the difference between two velocities (mask plus test and mask alone); it cannot report merely the direction of movement.

D. Application of the Gain-Control Model to Two-Flash Presentations

1. *An Asymmetric Saturation Theory*

Nakayama and Silverman[20] measured motion-direction discrimination in two-flash displays of sine gratings that translated θ degrees between flashes. As θ departed from 90 deg, observers could compensate the loss of discriminability by increasing the modulation amplitude of the gratings. However, for θ sufficiently near 0 or near 180 deg, no increase in amplitude was sufficient. These observations led Nakayama and Silverman to a theory of motion saturation to account for their results. Let m_1 represent the modulation amplitude of the first grating and m_2 the modulation amplitude of the second grating.

Their theory states that, in the absence of saturation, the output y of the motion detector is

$$y = \alpha m_2 \sin(\theta), \tag{8}$$

where α is a positive constant. If for a small θ the threshold modulation amplitude m_2 exceeds what is predicted by Eq. 8, the loss is ascribed to contrast saturation. Accordingly they derived a motion contrast-saturation function that we designate $g_N(m_2) = m_{\mathrm{eff}}$, which gives the net effective contrast m_{eff} after saturation has occurred (after gain control in our terminology). The proposed saturation function $g_N(m)$ was $g_N(m) = m$ for $m \leqslant 0.04$ and $g_N(m) = 0.04$ for $m \geqslant 0.04$.

Nakayama and Silverman's theory gives excellent fits to their data, but it has a singular problem: In their experiments, m_1 and m_2 were always equal, so a theory in which only m_2 appears is acceptable. In general, however, such a theory would be absurd because both flashes contribute, and the amplitude m_1 of the first flash cannot be ignored. Unfortunately, their particular calculation is intrinsically asymmetrical with respect to the two flashes. It does not generalize and it does not represent a plausible physiological process.

2. *Gain Control Plus Reichardt Motion Detector*

The output of a Reichardt model for two-flash displays is proportional to $m_1 m_2 \sin(\theta)$ and is intrinsically completely symmetrical with respect to the two flashes.[41] When we take into account a contrast gain-control function g_1, the Reichardt model predicts an output to a double flash proportional to $g_1(m_1)g_1(m_2)\sin(\theta)$. For the special case of two equal-contrast flashes, $m_1 = m_2 = m$. At the point where the output is just at the threshold for 75% correct, we have for the two theories

$$\begin{array}{ccc} \text{Reichardt+Expt. 1} & & \text{Nakayama–Silverman} \\ v_{75} = \beta g_1(m) g_1(m) \sin(\theta) & = & \alpha G_N(M) \sin(\theta). \end{array} \tag{9}$$

In Eq. (9), positive constants α and β are needed to scale the output, because the choice of v_{75} is arbitrary. It is apparent immediately from Eq. (9) that, for the special case of two equal flashes, the two theories can be made equivalent if we choose $g_1(m) = G_N(m)^{0.5}$. However, because of its aberrant behavior for small m, $G_N(m)^{0.5}$ is not a reasonable choice of $g_1(m)$. The question that we pose is the following: Does the model of Fig. 7a, a Reichardt detector with the gain-control function measured in Experiment 1, predict the Nakayama–Silverman data? (See Fig. 8.)

Equations (7) describe the gain-control mechanism of contrast saturation. The parameters for Eqs. (7) derived from Experiment 1 (pedestal masking of motion) are $k' = 0.0025$, $\eta' = 1.01$, and $\lambda' = 0.35$, for subject ZL. Experiment 1 used a five-frame moving sinusoid with a spatial frequency of 1.26 c/deg. Nakayama and Silverman presented two frames of a 2-c/deg grating, and their temporal parameters were different. These differences in procedure and subjects resulted in a slightly higher threshold. In Eqs. (7) the slight overall increase in threshold is represented by a slight reduction of k' and λ' (multiplying k', λ' by 0.95, the square root of the sensitivity change). Using these modified values yields the predictions of the Nakayama–Silverman data shown in Fig. 8, which also shows the original data and theory.

With just one free parameter to adjust for the slightly different overall thresholds in the two experiments, the shunting-plus-Reichardt model accounts for the two-flash data even better than the original Nakayama–Silverman computation ($r^2 = 0.992$ in the Nakayama–Silverman model, $r^2 = 0.993$ in shunting-plus-Reichardt).

3. *Conclusions: Two-Flash Predictions*

Let $g_1(m)$ be the output modulation of the gain-control function of the shunting-plus-Reichardt model when the input modulation is m. Let $G_N(m)$ be the corresponding Nakayama–Silverman gain-control function. Then (1) choosing $g_1(m)$ so that $g_1(m) = G_N(m)^{0.5}$ makes the shunting-plus-Reichardt and the Nakayama–Silverman model computationally equivalent within the symmetric two-flash experiment for which the Nakayama–Silverman model was derived. (2a) The Nakayama–Silverman model cannot be generalized even to two flashes of different contrasts, and so it is rejected. (2b) The shunting-plus-Reichardt model is a fully computational process model and generalizes to any stimulus configuration, but the gain-control function $g_1(m) = G_N(m)^{0.5}$ is unreasonable, and so this form of $g_1(m)$ is rejected. (3) Using instead the gain-control function derived from the first-order motion pedestal experiment together with the Reichardt model yields fits of the Nakayama–Silverman two-flash data that are as good as any previous fits, even though the components and the parameters of this model (contrast gain control from Experiment 1, motion detection by means of the Reichardt model) were derived independently of the two-flash data. (4) The identical model accounts for both the pedestal masking data and the two-flash threshold data; a model differing only slightly in the parameter η' [of Eqs. (7)] accounts for a great variety of other motion-masking mask-

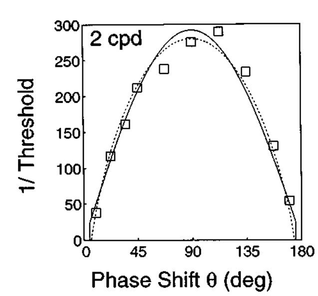

Fig. 8. Modulation thresholds for two-flash presentations of a sinusoidal grating as a function of the translation between flashes. Data are taken from Fig. 5A of Nakayama and Silverman[20] [2 cycles per degree (cpd) condition]. The horizontal axis denotes the phase angle θ between the two frames. The vertical axis denotes contrast sensitivity (1/threshold). Solid curve, prediction of Nakayama and Silverman [Eq. (8) in the text]; dotted curve, direct prediction based on the model of Fig. 7a (shunting feedforward gain control with a Reichardt motion detector) with the parameters taken directly from subject ZL in Experiment 1. One parameter was estimated from the data—a slight (0.95) sensitivity reduction in the two-flash (versus the present five-flash) experiment. Although both theories give statistically excellent predictions, only the shunting-plus-Reichardt theory generalizes to other paradigms (see text for details).

ing data. This universal applicability supports the original assumptions under which the model was derived: (1) gain control of motion signals occurs before the motion computation itself, (2) the mechanism of gain control is shunting feedforward gain control, and (3) the mechanism of motion detection is a Reichardt (or an equivalent motion-energy) computation.

E. Model for Motion Gain Control in the Second-Order System

In second-order motion, the motion signal is carried by a moving modulator imposed on a carrier texture. Experiment 2 showed that even a very-high-amplitude pedestal (a modulator of the same spatial frequency as the motion stimulus) failed to mask the motion stimulus. Experiment 3 showed that motion masking did occur when carrier amplitude was varied. It is astounding that the actual pattern being detected, the drifting sine grating, cannot be masked by a static sine grating of the same spatial frequency, but that, in fact, is a counterintuitive property of Reichardt (motion-energy) detectors. The failure of pedestals to mask held only for a limited range of pedestals in first-order stimuli because of prior gain control. In second-order motion detection, the prior gain control is not sensitive to the pedestal modulation, and therefore masking failure holds over the entire range of physically obtainable pedestals. Given these two facts (1) masking by a static carrier texture and (2) no masking by a static modulator, construction of a model of second-order motion gain control is quite straightforward.

Figure 7c shows a model of the second-order motion system. The carrier signal is carried by all the different frequency channels that happen to be stimulated by the carrier texture. Each channel is indicated as having a radially symmetrical center-surround field. Only three are shown, but there is assumed to be a continuum of such channels. In each local spatial area, the output of each channel is rectified, and the sum of all these rectified outputs is summed to compose the internal representation of the carrier (for the purposes of motion computation). This summed signal in the second-order system is now equivalent to a luminance signal in the first-order system. It has a direct path to the motion detector, which detects moving modulations that may be superimposed on it. This same carrier signal, spatially weighted to represent the neighborhood in which detection is occurring but without any further modification, also serves as the controlling signal in the gain-control pathway. The second-order system is structurally equivalent to the first-order system; the only differences are in how stimuli are preprocessed before gain control and motion detection.

ACKNOWLEDGMENT

This research was supported by the U.S. Air Force Office of Scientific Research, Life Sciences, Visual Information Processing Program.

Address correspondence to Prof. George Sperling, Department of Cognitive Sciences SSP, University of California, Irvine, California 92697. e-mail: sperling @uci.edu.

REFERENCES AND NOTES

1. D. C. Hood and M. A. Finkelstein, “Sensitivity to light,” in *Handbook of Perception and Human Performance*, K. R. Boff, L. Kaufman, and J. P. Thomas, eds. (Wiley, New York, 1986), Vol. 1, Chap. 5.
2. H. R. Blackwell, “Contrast thresholds of the human eye,” J. Opt. Soc. Am. **36**, 624–646 (1946).
3. E. G. Heinemann, “Simultaneous brightness induction as a function of inducing and test field luminances,” J. Exp. Psychol. **50**, 89–96 (1955).
4. G. Sperling and M. M. Sondhi, “Model for visual luminance discrimination and flicker detection,” J. Opt. Soc. Am. **58**, 1133–1145 (1968).
5. P. Whittle and P. D. C. Challands, “The effect of background luminance on the brightness of flashes,” Vision Res. **9**, 1095–1110 (1969).
6. G. Sperling, “Model of visual adaptation and contrast detection,” Percept. Psychophys. **8**, 143–157 (1970).
7. E. G. Heinemann, “Simultaneous brightness induction,” in *Handbook of Sensory Physiology*, D. Jameson and L. M. Hurvish, eds. (Springer Verlag, Berlin, 1972), Vol. VII/4, pp. 146–149.
8. R. Shapley and C. Enroth-Cugell, “Visual adaptation and retinal gain controls,” in *Progress in Retinal Research*, N. Osborne and G. Chader, eds. (Pergamon, Oxford, 1984), pp. 263–346.
9. G. Sperling, “Three stages and two systems of visual processing,” Spatial Vision **4**, 183–207 (1989).
10. A. M. Derrington and P. Lennie, “Spatial and temporal contrast sensitivities of neurons in lateral geniculate nucleus of macaque,” J. Physiol. (London) **357**, 219–240 (1984).
11. E. Kaplan and R. M. Shapley, “X and Y cells in the lateral geniculate nucleus of macaque monkeys,” J. Physiol. **330**, 125–143 (1982).
12. A. F. Dean, “The relationship between response amplitude and contrast for cat striate cortical neurons,” J. Physiol. (London) **318**, 413–427 (1981).
13. D. G. Albrecht and D. B. Hamilton, “Striate cortex of monkey and cat: contrast response function,” J. Neurophysiol. **48**, 217–237 (1982).
14. I. Ohzawa, G. Sclar, and R. D. Freeman, “Contrast gain control in the cat visual cortex,” Nature (London) **298**, 266–268 (1982).
15. G. Sclar, J. H. Maunsell, and P. Lennie, “Coding of image contrast in central visual pathways of the macaque monkey,” Vision Res. **30**, 1–10 (1990).
16. A. B. Bonds, “Temporal dynamics of contrast gain in single cells of the cat striate cortex,” Visual Neurosci. **6**, 239–255 (1991).
17. D. G. Albrecht and W. S. Geisler, “Motion selectivity and the contrast-response function of simple cells in the visual cortex,” Visual Neurosci. **7**, 531–546 (1991).
18. D. J. Heeger, “Normalization of cell responses in cat striate cortex,” Visual Neurosci. **9**, 181–197 (1992).
19. D. G. Pelli, “Effects of visual noise,” Ph.D. dissertation (University of Cambridge, Cambridge, 1980).
20. K. Nakayama and G. H. Silverman, “Detection and discrimination of sinusoidal grating displacements,” J. Opt. Soc. Am. A **2**, 267–274 (1985).
21. J. H. Jamar and J. J. Koenderink, “Contrast detection and detection of contrast modulation for noise gratings,” Vision Res. **25**, 511–521 (1985).
22. S. P. McKee, G. H. Silverman, and K. Nakayama, “Precise velocity discrimination despite random variations in temporal frequency and contrast,” Vision Res. **26**, 609–619 (1986).
23. M. Pavel, G. Sperling, T. Riedl, and A. Vanderbeek, “Limits of visual communication: the effect of signal-to-noise ratio on the intelligibility of American Sign Language,” J. Opt. Soc. Am. A **4**, 2355–2365 (1987).
24. D. H. Parish and G. Sperling, “Object spatial frequencies, retinal spatial frequencies, noise, and the efficiency of letter discrimination,” Vision Res. **31**, 1399–1415 (1991).
25. J. G. Robson, “Spatial and temporal contrast-sensitivity functions of the visual system,” J. Opt. Soc. Am. **56**, 1141–

1142 (1966).
26. D. H. Kelly, "Adaptation effects on spatio-temporal sine-wave threshold surface," Vision Res. **12**, 89–101 (1972).
27. D. H. Kelly, "Motion and vision. II. Stabilized spatio-temporal threshold surface," J. Opt. Soc. Am. **69**, 1340–1349 (1979).
28. J. J. Koenderink and A. J. van Doorn, "Spatiotemporal contrast detection threshold surface is bimodal," Opt. Lett. **4**, 32–34 (1979).
29. D. C. Burr and J. Ross, "Contrast sensitivity at high velocities," Vision Res. **22**, 479–484 (1982).
30. J. P. H. van Santen and G. Sperling, "Temporal covariance model of human motion perception," J. Opt. Soc. Am. A **1**, 451–473 (1984).
31. Z.-L. Lu and G. Sperling, "The functional architecture of human visual motion perception," Vision Res. **35**, 2697–2722 (1995).
32. J. Nachmias and R. V. Sansbury, "Grating contrast: discrimination may be better than detection," Vision Res. **14**, 1039–1042 (1974).
33. C. F. Stromeyer and S. Klein, "Spatial frequency channels in human vision as asymmetric (edge) mechanisms," Vision Res. **14**, 1409–1420 (1974).
34. G. E. Legge and J. M. Foley, "Contrast masking in human vision," J. Opt. Soc. Am. **70**, 1458–1471 (1980).
35. C. A. Burbeck and D. H. Kelly, "Contrast gain measurements and the transient/sustained dichotomy," J. Opt. Soc. Am. **71**, 1335–1342 (1981).
36. M. Livingstone and D. Hubel, "Segregation of form, color, movement and depth: anatomy, physiology, and perception," Science **240**, 740–749 (1988).
37. S. J. Anderson, D. C. Burr, and M. C. Morrone, "Two-dimensional spatial and spatial-frequency selectivity of motion-sensitive mechanisms in human vision," J. Opt. Soc. Am. A **8**, 1340–1351 (1991).
38. A. T. Smith, "Correspondence-based and energy-based detection of second-order motion in human vision," J. Opt. Soc. Am. A **11**, 1940–1948 (1994).
39. W. Reichardt, "Autokorrelationsauswertung als Funktionsprinzip des Zentralnervensystems," Z. Naturforsch. **12b**, 447–457 (1957).
40. W. Reichardt, "Autocorrelation, a principle for the evaluation of sensory information by the central nervous system," in *Sensory Communication*, W. A. Rosenblith, ed. (Wiley, New York, 1961), pp. 303–317.
41. J. P. H. van Santen and G. Sperling, "Elaborated Reichardt detectors," J. Opt. Soc. Am. A **2**, 300–321 (1985).
42. E. H. Adelson and J. K. Bergen, "Spatio-temporal energy models for the perception of apparent motion," J. Opt. Soc. Am. A **2**, 284–299 (1985).
43. A. B. Watson and A. J. Ahumada, Jr., "A look at motion in the frequency domain," in *Motion: Perception and Representation*, J. K. Tsotos, ed. (Association for Computing Machinery, New York, 1983), pp. 1–10.
44. P. Cavanagh and G. Mather, "Motion: the long and the short of it," Spatial Vision **4**, 103–129 (1989).
45. C. Chubb and G. Sperling, "Two motion perception mechanisms revealed by distance driven reversal of apparent motion," Proc. Natl. Acad. Sci. (USA) **86**, 2985–2989 (1989).
46. D. J. Heeger, "A model for the extraction of image flow," J. Opt. Soc. Am. A **4**, 1455–1471 (1987).
47. H. R. Wilson, V. P. Ferrera, and C. Yo, "A psychophysically motivated model for two-dimensional motion perception," Visual Neurosci. **9**, 79–97 (1992).
48. S. J. Nowlan and T. J. Sejnowski, "Filter selection model for motion segmentation and velocity integration," J. Opt. Soc. Am. A **11**, 3177–3200 (1994).
49. C. Chubb and G. Sperling, "Drift-balanced random stimuli: a general basis for studying non-Fourier motion perception," J. Opt. Soc. Am. A **5**, 1986–2006 (1988).
50. C. Chubb and G. Sperling, "Texture quilts: basic tools for studying motion-from-texture," J. Math. Psychol. **35**, 411–442 (1991).
51. V. S. Ramachandran, M. V. Rau, and T. R. Vidyasagar, "Apparent movement with subjective contours," Vision Res. **13**, 1399–1401 (1973).
52. A. M. M. Lelkens and J. J. Koenderink, "Illusory motion in visual displays," Vision Res. **24**, 293–300 (1984).
53. A. M. Derrington and D. R. Badcock, "Separate detectors for simple and complex grating patterns?" Vision Res. **25**, 1869–1878 (1985).
54. K. Turano and A. Pantle, "On the mechanism that encodes the movement of contrast variations—I: velocity discrimination," Vision Res. **29**, 207–221 (1989).
55. J. D. Victor and M. M. Conte, "Motion mechanisms have only limited access to form information," Vision Res. **30**, 289–301 (1989).
56. D. Marr and S. Ullman, "Directional selectivity and its use in early visual processing," Proc. R. Soc. London Ser. B **211**, 151–180 (1981).
57. M. S. Landy, Y. Cohen, and G. Sperling, "HIPS: a Unix-based image processing system," Comput. Vision Graphics Image Process. **25**, 331–347 (1984a).
58. M. S. Landy, Y. Cohen, and G. Sperling, "HIPS: image processing under Unix-software and applications," Behav. Res. Methods Instrum. Comput. **16**, 199–216 (1984b).
59. Runtime Library for Psychology Experiments, Human Information Processing Laboratory, Department of Psychology, New York University, New York 10003, 1988.
60. R. S. Woodworth and H. Schlosberg, *Experimental Psychology*, rev. ed. (Holt, Rinehart & Winston, New York, 1954).
61. K. I. Naka and W. A. Rushton, "S-potentials from colour units in the retina of fish (*Cyprinidae*)," J. Physiol. (London) **185**, 536–555 (1966).
62. J. S. Coombs, J. C. Eccles, and P. Fatt, "The specific ionic conductances and ionic movements across the motoneuronal membrane that produce the inhibitory post-synaptic potential," J. Physiol. **130**, 326–373 (1955).
63. A. B. Watson, P. G. Thompson, B. J. Murphy, and J. Nachmias, "Summation and discrimination of gratings moving in opposite directions," Vision Res. **20**, 341–347 (1980).
64. D. H. Kelly, "Motion and vision. IV. Isotropic and anisotropic spatial responses," J. Opt. Soc. Am. **72**, 432–439 (1982).
65. C. R. Carlson and R. W. Klopfenstein, "Spatial-frequency model for hyperacuity," J. Opt. Soc. Am. A **2**, 1747–1751 (1985).
66. S. A. Klein and D. M. Levi, "Hyperacuity threshold of 1 sec: theoretical predictions and empirical validation," J. Opt. Soc. Am. A **2**, 1171–1190 (1985).