PERGAMON

Vision Research 41 (2001) 2355–2374

Vision Research

www.elsevier.com/locate/visres

Sensitive calibration and measurement procedures based on the amplification principle in motion perception

Zhong-Lin Lu [a,*], George Sperling [b]

[a] *Department of Psychology, University of Southern California, Los Angeles, CA 90089-1061, USA*
[b] *Departments of Cognitive Sciences and of Neurobiology and Behavior, and the Institute of Mathematical Behavioral Sciences, University of California, Irvine, CA 92697, USA*

Received 4 March 1999; received in revised form 3 January 2001

Abstract

We compare two types of sampled motion stimuli: ordinary periodic displays with modulation amplitude $m_{o=e}$ that translate 90° between successive frames and amplifier sandwich displays. In sandwich displays, even-numbered frames are of one type, odd-numbered frames are of the same or different type, and (1) both types have the same period, (2) translate in a consistent direction 90° between frames, and (3) even frames have modulation amplitude m_e, odd frames have modulation amplitude m_o. In both first-order motion (van Santen, J.P.H. & Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America A*, *1*, 451–73) and second-order motion (Werkhoven, P., Sperling, G., & Chubb, C. (1993). Motion perception between dissimilar gratings: a single channel theory. *Vision Research*, *33*, 463–85) the motion strength of amplifier sandwich displays is proportional to the product $m_o m_e$ for a wide range of m_e. By setting m_e to a large value, an amplifier sandwich stimulus with a very small value of m_o can still produce visible motion. The amplification factor is the ratio of two threshold modulation amplitudes: ordinary $\hat{m}_{o=e}$ over amplified $\hat{m}_o$, $\hat{m}_{o=e}/\hat{m}_o$. We find amplification factors of up to about $8\times$. Light adaptation and contrast gain control in early visual processing distort the representations of visual stimuli so that inputs to subsequent perceptual processes contain undesired distortion products or 'impurities'. Motion amplification is used to measure and thence to reduce these unwanted components in a stimulus to a small fraction of their threshold. Such stimuli are certifiably pure in the sense that the residual impurity is less than a specified value. Six applications are considered: (1) removing (first-order) luminance contamination from moving (second-order) texture gratings; (2) removing luminance contamination from moving chromatic gratings to produce pure isoluminant gratings; (3) removing distortion products in luminance-modulated (first-order) gratings — by iterative application, all significant distortion products can be removed; (4) removing second-order texture contamination from third-order motion displays; (5) removing feature bias from third-order motion displays; (6) and the same general principles are applied to texture-slant discrimination in which x, y spatial coordinates replace the x, t motion coordinates. In all applicable domains, the amplification principle provides a powerful assay method for the precise measurement of very weak stimuli, and thereby a means of producing visual displays of certifiable purity.

Keywords: Amplification; Calibration; First-order motion; Second-order motion; Third-order motion; Reichardt model; Motion-energy; Measurement

1. Introduction

We analyze here a method for measuring weak stimuli or, more precisely, weak stimulus components. When stimuli are processed by early stages of the visual system, they undergo light adaptation, contrast gain control, and other nonlinear processes. These early transformations distort even low-contrast stimuli, so that subsequent visual processes receive an input that can be regarded as having acquired undesired distortion products or 'impurities'. Even when there is no distortion, a stimulus may unintentionally stimulate more than one system. For example, a nominally isoluminant red–green grating may stimulate a luminance system.

* Corresponding author. Tel.: +1-213-7402282; fax: +1-213-7469082.
E-mail addresses: zhonglin@rcf.usc.edu (Z.-L. Lu), sperling@uci.edu (G. Sperling).

0042-6989/01/$ - see front matter
PII: S0042-6989(01)00106-7

Our aim is to describe a paradigm for reducing such undesired impurities to far below their threshold of visibility. Mainly, motion stimuli are considered, although the method is quite general. It is illustrated in the context of developing motion displays that stimulate only one particular motion system or only one motion computation (within a motion system). We begin with a definition of what is meant here by a 'motion system'.

1.1. Background: motion systems

Recent psychophysical experiments have led us to propose a three-systems functional architecture for human visual motion perception (Lu & Sperling, 1995a; Sperling & Lu, 1998a). A motion system is a computation, presumably carried out in a brain nucleus, that takes an input which is a function of space x, y and time t, and computes as its output a motion flowfield $\mathbf{f}(x, y, t)$. $\mathbf{f}(x, y, t)$ is a vector that represents the magnitude and direction of motion at the point x, y at time t.

The first-order motion system computes flowfields from moving luminance modulations by means of a primitive motion-energy (or equivalently, Reichardt detector) algorithm (Reichardt, 1957; Watson & Ahumada, 1983; van Santen & Sperling, 1984, 1985; Adelson & Bergen, 1985).

The second-order system (originally called 'non-Fourier motion' by Chubb & Sperling, 1989a,b) extracts motion from stimuli with moving feature modulations in which the expected luminance is the same everywhere but in which some features (e.g. texture contrast) move systematically in space and time. The second-order system employs a texture-grabber (spatiotemporal linear filtering plus fullwave rectification, see Chubb & Sperling, 1989a,b) to compute the amount of features, and then applies the same motion energy algorithm to features as the first-order system does to photons (Lu & Sperling, 1995a; Sperling & Lu, 1998a).

The third-order system detects movement of 'salience'. It is useful to regard salience as recorded in a salience map $s(x, y, t)$ in which the x, y, t locations of areas indicated as 'figure' (vs. 'ground') are marked with positive values (e.g. $s(x, y, t) = 1$) whereas locations indicated as ground or as unimportant are assigned zero (Lu & Sperling, 1995b; Blaser, Sperling, & Lu, 1999). Figure–ground is a binary concept. Not all areas can be unambiguously classified, and even among similarly classified areas, there may be differences in importance. Salience is a continuous variable that incorporates both 'figureness' and 'importance'. The third-order motion system is assumed to compute motion from the dynamic salience map, i.e. the motion of important locations as a function of time.

Whereas these three motion systems carry out parallel and independent computations, a particular motion display typically stimulates more than one of the three systems (Lu & Sperling, 1995a, 1996a; Sperling & Lu, 1998a; Lu & Sperling, 2001). For example, a drifting sinusoidal luminance grating stimulates all three systems. To better reveal the properties of individual motion systems, it is desirable to produce motion displays that stimulate only one motion system. Here we address the question of how to generate certifiably 'pure' motion displays. Recent reports (Smith & Ledgeway, 1997; Sperling & Lu, 1998b; Lu & Sperling, 2001; Scott-Samuel & Georgeson, 1999) of luminance (first-order) contamination in second-order motion stimuli make this a critical research issue.

1.2. Sandwich displays and multiplicative amplification

To guide the discussion, we consider a specific example in which the aim is to eliminate a possible first-order (luminance motion) component from a second-order (contrast-modulated texture) motion display. The contrast-modulated texture display consists of a sequence of texture gratings with a squarewave contrast modulation (Fig. 1a and b). In such a random texture grating, the expected luminance is the same everywhere, and indeed, average luminance over even small areas is nearly the same and contains no systematic luminance variation. Successive frames are identical to the first except that the contrast modulation is translated by 90° from frame to frame. Typically, such a display produces vivid apparent motion. To determine whether this motion might be caused by unintentional stimulation of a first-order (luminance) motion system (presumably caused by early perceptual distortion of the intensities in the display), there are two useful methods: a sandwich paradigm[1] that utilizes the motion-amplification principle, and a minimum motion method.[2] The ultimate aim of both methods is to add to the stimulus frames a corrective modulation that cancels the presumed perceptual distortion products and thereby produces a pure (in this case, pure second-order) stimulus. We consider motion-amplification first.

[1] Early uses of sandwich displays were reducing luminance contamination in red–green chromatic motion (Anstis & Cavanagh, 1983), and demonstrating the multiplicative property of motion perception predicted by the elaborated Reichardt detector (van Santen & Sperling, 1984). Sandwich displays to estimate perceptually produced luminance contamination in contrast-modulated texture motion were used by Scott-Samuel and Georgeson (1999) and by Sperling and Lu (1998b).

[2] An early use of the minimum motion method (to reduce luminance contamination in red–green chromatic motion) is Moreland (1980).

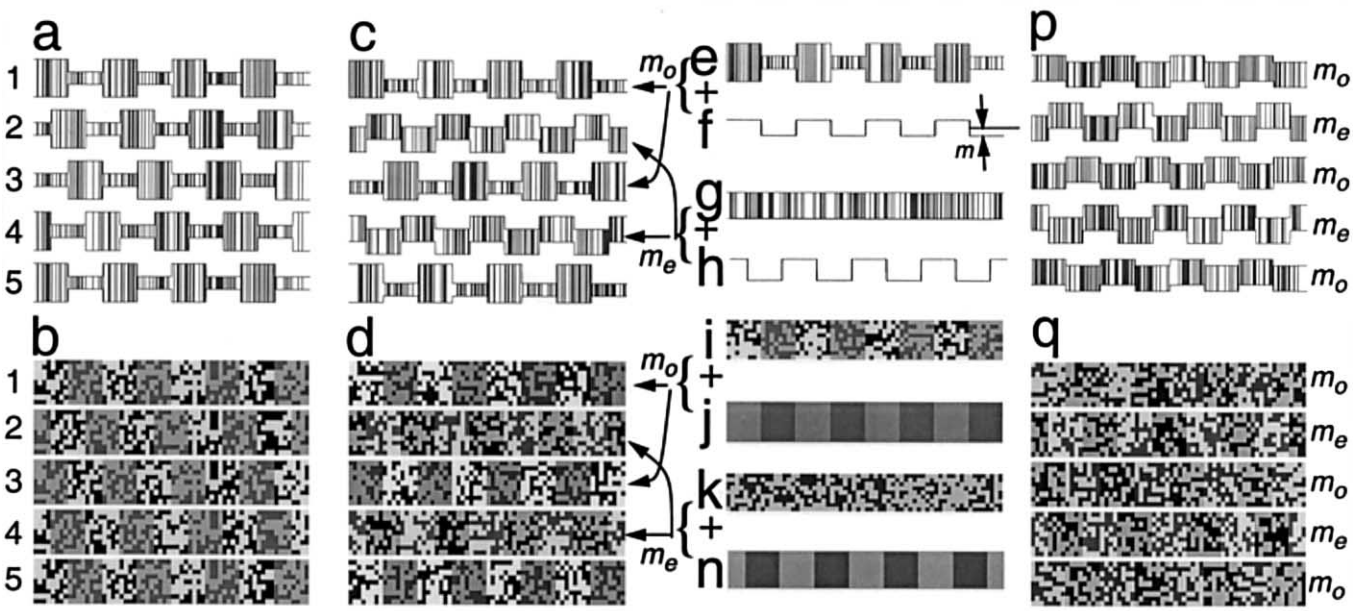

Fig. 1. Displays for measuring first-order (luminance) components in second-order (texture-contrast) motion. The panels show five frames of a moving sinewave grating. The upper row illustrates the luminance in one horizontal slice of each successive frame; the lower row illustrates the appearance of the frames themselves. The phase shift between successive frames is 90° in all instances. (a, b) Five frames of a contrast-modulated texture grating in which the modulation translates 90° between successive frames. The carrier (the texture pattern) is 'static' — it does not change between frames. In (a), the abscissa represents space, x; the ordinate represents intensity of one row of the frame. Expected luminance is constant throughout the entire stimulus. (c–n) Composition of an amplification sandwich display (c). The odd frames (1, 3, 5) of (c) are similar to the odd frames of (a). They are composed of a carrier which is *multiplied* by a modulator (g). A small luminance component (f) is added to (e) to cancel distortion components produced in early vision. In even frames (2, 4), a large luminance modulation (h) is *added* to the texture carrier (g) to produce luminance amplification. (i, j, k, n) represent image slices corresponding to (e, f, g, h). (p, q) A sandwich display for measuring amplification of luminance motion in the presence of a random texture carrier. Each frame is composed of a uniform texture (g, k) to which is added a luminance modulation (h, n). The carrier is 'dynamic' — each frame is different. The modulation depth in even frames m_e is set to a fixed value; the modulation depth in the odd frames m_o is varied to determine the motion-direction threshold $\acute{m}_o$.

1.2.1. The amplifier sandwich display

In an amplifier sandwich (motion) display, two types of frames, test frames and amplifier frames alternate.[3] In the texture-motion example, the test frames are contrast-modulated texture gratings. For reasons that will be considered below, usually the odd frames contain the test gratings. Frame 1 of the sandwich display consists of any frame selected from the contrast-modulated texture motion sequence (Fig. 1a). Frame 3 is identical to frame 1 except that the contrast modulation is shifted by 180°, i.e. the high-contrast strips fall where the low-contrast strips occurred previously. Frame 5 is again identical to the first frame, and so on. Because successive odd frames translate by 180°, the odd frames are merely a counterphase modulation of the original frame and, by themselves, are completely ambiguous with respect to motion direction (Fig. 1c and d).

In an amplifier sandwich display, even frames contain precisely the same carrier texture as the odd frames, but there is zero contrast modulation. However, a luminance modulation (Fig. 1g,h,k,n) that has the same spatial frequency as the contrast modulation in the odd frames is added to the even frames. Successive even frames also differ by 180° in phase. Therefore, there is neither a motion stimulus within the even frames nor within the odd frames. Only a common feature in both, e.g. the possible perceptually-generated luminance contamination in the texture frames and the real luminance contamination can produce apparent motion. The even frames will function as amplifier frames to sensitively reveal perceptually produced luminance modulation, if any, in the odd frames.

Suppose sufficient luminance modulation were to be physically added to the odd frames of the sandwich display (Fig. 1e,f,i,j). Of course, luminance motion would then be perceived in the sandwich display. The sandwich display would then be equivalent to a luminance motion display, with two unequal amplitudes of luminance motion, superimposed on a texture pattern, which functions as a masking pattern. The perceived motion direction (left or right) would depend on the phase ($+$ or $-$) of the added luminance modulation. When the phase and amplitude of a luminance modula-

[3] 'Sandwich system' is sometimes used to refer to a system in which a nonlinear stage is sandwiched between two linear stages. Here the use refers strictly to motion displays in which test frames (bread) and amplifier frames (meat) alternate.

tion added to the test (odd) frames is such that the luminance contamination is precisely cancelled, a point of minimum motion (or maximum motion-direction ambiguity) is reached. It is assumed that, at this point, the residual contamination is at a minimum (Fig. 2b).

1.2.2. Minimum motion method

In the minimum motion method, the modulation is translated by a constant amount in a consistent direction from frame to frame (e.g. Fig. 1a–b and p–q). In chromatic motion, e.g. of a red–green grating, this means translating the image itself. In second-order motion, for example, in contrast-modulated texture motion, the contrast modulator is translated (Fig. 1a) while the carrier texture remains stationary. A particular amplitude m of luminance modulation is added in a particular phase θ to the translating modulation (e.g. Fig. 1e,f,i,j). A search is then made for the m, θ combination that produces minimum motion. Usually, only one phase is tested, so the search is for the value of m that produces minimum motion (Fig. 2b). This added m, θ combination is assumed to cancel the contamination component.

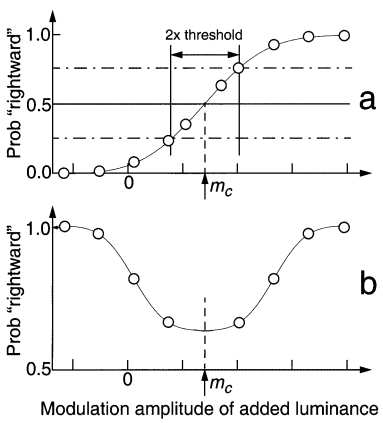

Fig. 2. Motion-direction judgments as a function of the magnitude of an added 'correction' component in (a) amplification sandwich and (b) minimum motion displays. (a) Psychometric function, $P_c(m_o)$. The abscissa is the modulation amplitude m_o of a 'correction' component. For example, it is the modulation amplitude m_o of a luminance (first-order) grating added to the odd frames of a sandwich display, which contain a contrast-modulated texture grating (e.g. Fig. 1c,d), while the even (amplifier) frames have a luminance modulation m_e. The ordinate is the estimated probability that the judged motion direction is consistent with the nominal stimulus motion direction. The point of minimum motion (maximum motion ambiguity) is $P_c(m_c) = 50\%$; at this point the first-order component in the odd frames is cancelled exactly. (b) Idealized data from a minimum motion procedure. The motion direction varies randomly from trial to trial. The correction component is added identically to every frame. The ordinate is probability of correctly perceiving the direction of motion; the abscissa is as in (a). In the hypothetical data of (b), the added correction (e.g. a luminance grating in the case of a moving contrast modulation) cancels the luminance contamination but some residual ability to discriminate motion survives.

In the minimum motion method it is absolutely critical that the two motions, the possible contamination and the added cancelling modulation, both move in the same direction. The ability to cancel a motion component by another motion component moving in the same direction is a critical test of both motion components activating the same motion detector(s) (see Lu & Sperling, 1995a). When motion in one direction is cancelled by a motion component in the opposite direction, the cancellation may occur at a higher level, after two independent motion computations (e.g. Chichilnisky, Heeger, & Wandell, 1993).

1.2.3. Multiplicative amplification, m_o m_e

Reichardt (1961), in exploring the motion perception of beetles, observed that when adjacent facets of their compound eye were stimulated successively with two positive flashes of light, $+1$, $+1$, or two dark flashes, -1, -1, motion was perceived in one direction, but the sequences $+1$, -1, or -1, $+1$, produced motion in the reverse direction (a phenomenon rediscovered and named 'reverse phi' by Anstis, 1970). On this basis, Reichardt proposed his correlation model (actually a covariance model, van Santen & Sperling, 1984) for beetle motion-direction perception. van Santen and Sperling (1984) elaborated the Reichardt model for human vision (the elaborated Reichardt detector, ERD). They derived and tested the multiplicative property of the ERD. Obviously, the multiplicative property also holds for the motion energy model (Adelson & Bergen, 1985) and other models that have been shown to be equivalent or nearly equivalent to the ERD (Watson & Ahumada, 1983; Adelson & Bergen, 1986).

When successive frames are identical in waveform except for amplitude, and when the frame-to-frame translation is 90°, the strength of motion is proportional to $m_o m_e$, the modulation amplitude of the odd frames times the modulation amplitude of the even frames (van Santen & Sperling, 1984, Eq. (19), p. 456). A sinewave amplifier sandwich display that was used to test the multiplicative property is illustrated in Fig. 3, but the multiplicative property of Reichardt/motion-energy models holds for any waveform. Indeed, the multiplicative property holds for any ERD, no matter what its parameters may be, no matter where it may be placed relative to the stimulus, and no matter how many frames there may be in the stimulus. The multiplication principle can be used to arrive at a measurement scale for amplifier sandwich displays, and it suggests how to make the most sensitive scale; i.e. to gain maximum amplification from the sandwich method.

The multiplication property was empirically tested by van Santen and Sperling (1984) for 16 pairs of luminance-modulated (first-order) sinewave gratings (Fig. 3). Odd-frame modulations m_o were 2.3, 4.2, 6.4, 8.5%;

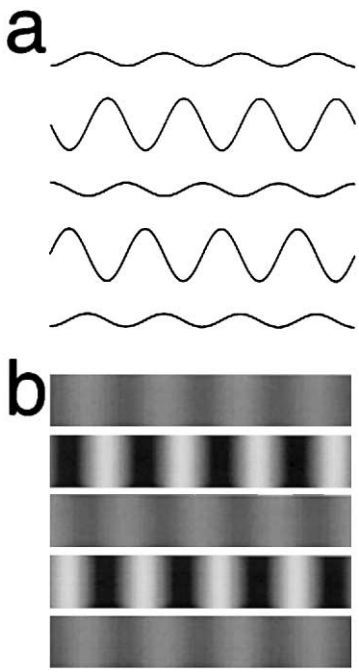

Fig. 3. A sandwich display for testing the amplification principle in first-order motion. Each panel shows five frames of a rightward moving sinewave grating. The upper row illustrates the luminance in one horizontal slice of each successive frame; the lower row illustrates the appearance of the frames themselves. The phase shift between successive frames is 90°. Odd frames have amplitude m_o; even frames, m_e.

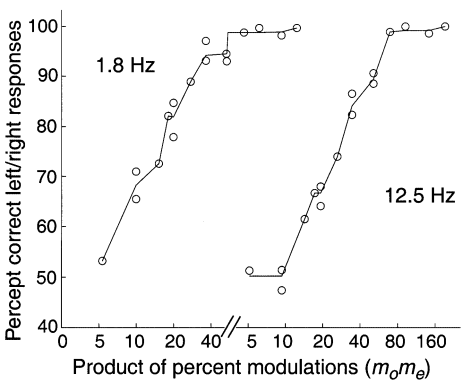

Fig. 4. Data from a sandwich experiment to test the multiplicative property for first-order sinewave motion. Data are from one observer, two temporal frequencies (1.8, 12.5 Hz), and 16 different $m_o m_e$ combinations (Fig. 3). The ordinate is the accuracy of motion-direction judgments, the abscissa is motion power $m_o m_e$ (van Santen & Sperling, 1984).

even-frame modulations m_e were 2.3, 4.2, 8.5, 22.9%. The probability P_c of a correct motion-direction response was a monotonic function simply of the product $m_e m_o$ (Fig. 4).

Different combinations of $m_e m_o$ that have similar products produce similar response accuracies P_c. This is true for temporal frequencies of 1.8 Hz (quite slow) and 12.5 Hz (fast). Results for another observer are similar except that accuracy with displays containing the largest modulation $m_o = 22.9\%$ is somewhat smaller than would be expected from the product relation.

In Fig. 4, the product relation holds for the entire range tested, from $P_c = 50\%$ (chance accuracy) to near perfect performance $P_c \approx 100\%$, approximately an 8:1 range of products. Consider two examples: (1) $m_e m_o = 4.2 \times 4.2\% = 17.6$ and $m_e m_o = 2.3 \times 8.5\% = 8.5 \times 2.3\% = 19.6$. At both temporal frequencies, all three stimuli produce the same response accuracy within measurement error: $P_c = 80\%$ (1.8 Hz), $P_c = 65\%$ (12.5 Hz). (2) $m_e m_o = 2.3 \times 22.9\% = 52.7$ and $m_e m_o = 6.4 \times 8.5 = 54.4$ both produce the same response accuracy within measurement error 92% (1.8 Hz), 89% (12.5 Hz).

The one reported counter example to the product rule for small- and moderate-amplitude (nonsaturating) first-order motion stimuli known to us (Nakayama & Silverman, 1985), was subsequently shown to fit the product relation perfectly when contrast gain control was considered. The product relation accounted for 98% of the variance of the N–S data and offered an easy generalization to other stimuli, which the original N–S theory did not (Lu & Sperling, 1996b).

1.2.4. The amplification factor

To illustrate the amplification factor, consider the stimuli of example 2. For the odd-frame stimulus, $m_o = 2.3\%$. A stimulus with both $m_e = m_o = 2.3\%$ produces only chance performance ($P_c = 50\%$) at either frequency. But by increasing m_e by a factor of 10 (to 22.9%) in combination with $m_o = 2.3\%$, response accuracy is increased to $P_c = 90\%$. Because the multiplicative property holds, to produce the same response accuracy in a simple sampled-motion stimulus in which all frames have the same modulation amplitude would require the same product of $m_e m_o$: $m_e = m_o = 2.3\sqrt{10} \approx 7.27\%$. Because substituting $m_o = 10 \times 2.3$ for 2.3 in the even frames of an amplifier sandwich produces the same motion strength as a sampled-motion stimulus with a modulation of about 7%, one might regard the sandwich as amplifying the odd frames by $3\times$.

The multiplication property fails for large amplifier modulations m_e. Because the multiplication property may fail, it is better to define an amplification factor that is independent of any theory, simply the ratio of two measured thresholds: the ordinary threshold divided by the sensitized threshold. Ordinary threshold, $\hat{m}_{o=e}$, is a threshold determined when $m_o = m_e$. Sensitized threshold $\hat{m}_o$, is a threshold for m_o determined when $m_e = k\hat{m}_{o=e}$, $k > 1$. When the multiplicative property holds, and when $m_e = km_{o=e}$, then $\hat{m}_o = (1/k)\hat{m}_{o=e}$. Motion amplification is simply k.

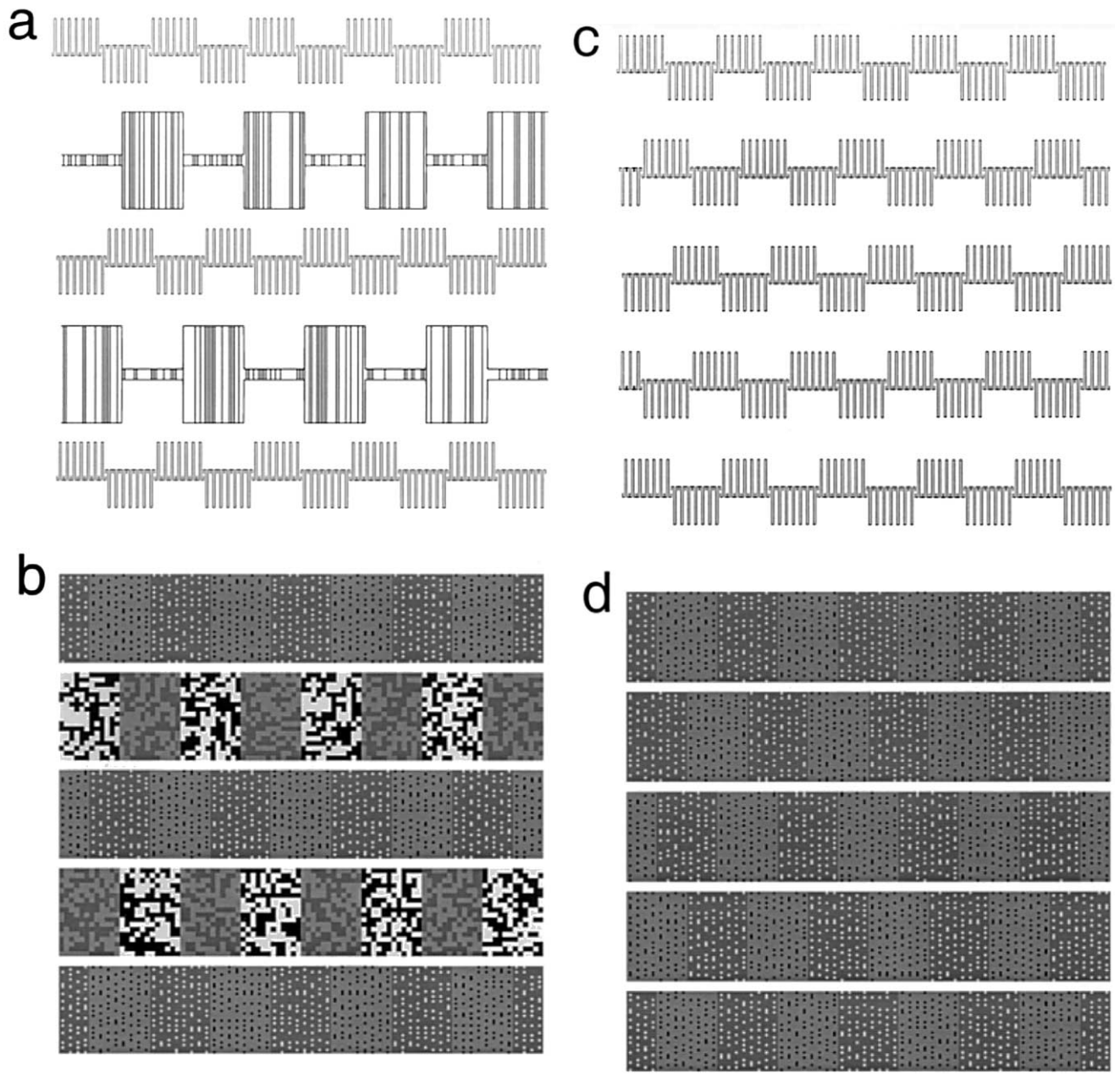

Fig. 5. Procedures for removing second-order fullwave contamination from presumptive third-order halfwave displays. Halfwave displays are composed of 'Mexican hats' — a micro-pattern consisting of 3×3 pixels in which the middle pixel has a high contrast c and the other eight pixels have a contrast $-c/8$ so that the mean luminance of the whole micro-pattern is 0. c is positive for white hats and negative for black hats. Initially, luminance (first-order) contamination is removed. Then, when fullwave (second-order) contamination is removed from the hat frames, the difference between black and white hats with the same amplitude can only be detected by a halfwave (third-order) system (Solomon & Sperling, 1994). (a, b) Odd frames: halfwave gratings made of alternating patches of black and white hats with 180° phase shift between successive odd frames. Even frames: High amplitude ($= 10 \times$ threshold) fullwave (second-order) texture gratings, with 180° phase shift between successive even frames. When there is fullwave contamination in the halfwave gratings, it, together with the fullwave gratings in the even frames produces consistent apparent motion. With the white hat contrast fixed, black hat contrast is varied until no consistent motion is perceived. (c, d) The minimum motion procedure as used by Solomon and Sperling (1994). Five frames of halfwave gratings are presented with 90° phase shift between successive frames. To keep the experiment in a measurable range, relatively low contrasts were used so that observers' performances in judging motion direction went from just above 50% (chance) at the curve minimum to almost 100% correct. Keeping the white contrast constant, contrast of the black hats was varied to minimize the accuracy of motion direction judgments. At the minimum performance level, the display was considered as clean (no fullwave contamination).

1.2.5. Multiplicative property of second-order motion

The multiplicative property was observed in second-order motion perception by Werkhoven, Sperling, and Chubb (1993). They used a motion path-competition paradigm. The motion strength of a path was determined by the multiplication of the activity from the even and the odd frames A_eA_o of texture patches along the path. The activities themselves, A_e and A_o, are the rectified outputs of a low-pass spatio-temporal filter — a texture grabber in the terminology of Chubb and Sperling, (1989a,b) — that extracts texture contrast. From Werkhoven et al. (1993), it can be inferred that, in second-order stimuli composed of texture gratings in which even and odd frames have different activity (contrast) modulations (Fig. 5a and b), a large m_e in amplifier frames could amplify the effect of a small, even 'invisible' m_o in test frames. This is the amplification principle in second-order motion perception. The second-order analog of first-order motion amplification is shown in Fig. 1e–h.

1.2.6. Two phases to test

To exactly cancel an undesired motion component (e.g. in the odd frames), an equal and oppositely signed component must added to the odd frames. For a sine grating, it seems reasonable that any contamination component would be exactly in-phase or exactly out-of-phase with the primary modulation. Indeed this would be true for any stimulus composed entirely of in-phase odd harmonics.

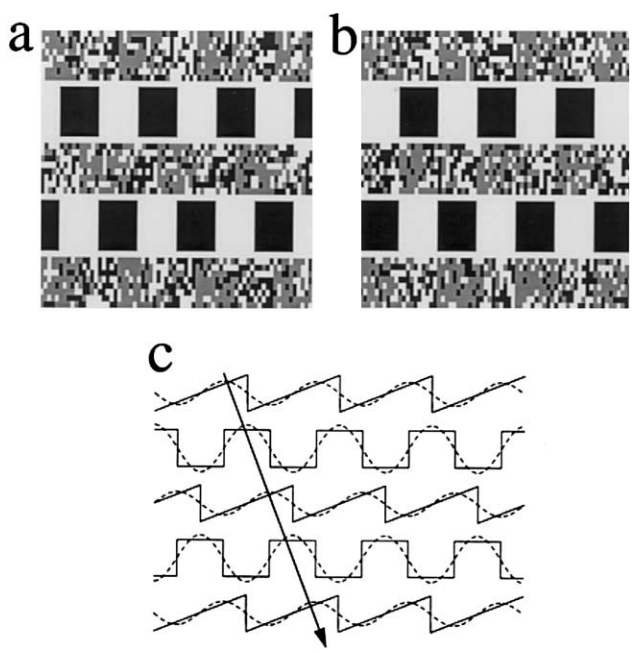

Fig. 6. Sandwich displays with 0 and 90° frame-to-frame phase shifts. (a) 90°, (b) 0° phase shift between successive frames. The density of texture increases in a sawtooth fashion, illustrated in (c). (c) Diagrammatic representation of the stimuli in (a). Physical sine components are illustrated in dotted curves. For the odd frames, the sawtooth illustrate texture density; there is no physical luminance sine component. However, early visual distortion produces a psychophysical sine components in the odd frames which does not necessarily correspond precisely in phase to the texture density, and does not necessarily have a 90° phase shift relative the even frames.

For a complex periodic stimulus that contains harmonics, it may not always be obvious what the exact phase of motion components might be. If there were contamination at any phase other than 0 or 180°, then an undesired motion component could be cancelled in the amplifier sandwich display (nominal 90° phase shifts between successive stimuli) but a residue would remain to be revealed in other displays. This residue might well be above motion threshold when the original display was simply translated. To be sure that this cannot occur, recall that any sinewave can be constructed from two sinewaves of the same frequency that differ in phase by 90°. So, sandwiches must be constructed not only with the even amplifier frames shifted 90° relative to odd test frames but also for unshifted even frames with 0° shifts relative to the preceding odd frame (Fig. 6a). Cancelling components are then added at both 90 and at 0° (with pure sine waves, this is equivalent to adding a sinewave at a particular phase) to cancel the undesired component. When the phase of a contamination component is unknown, it would be advisable to use a minimum motion paradigm to confirm the cancellation achieved with the amplifier sandwich method.

Stromeyer, Kronauer, Ryu, Chaparro, and Eskew (1995) demonstrate that different wavelength components have different perceptual latencies, as do Cavanagh and Anstis (1991).[4] In motion, a difference in time is equivalent to a difference in space. So, the method of calibrating with latency differences is the same as for calibrating with spatial asymmetry as described above: calibration with 0° as well as with 90° amplifier frames. This treats the latency difference in terms of an equivalent spatial difference. Obviously, such a calibration is useful only under identical temporal conditions as those used for the calibration.

In our experience with sinewave and squarewave gratings, we have never observed motion in amplifier sandwich displays with zero-shift even frames. However, a sawtooth waveform is severely asymmetric, and would require titration in both 90 and 0° sandwiches to cancel all phases of an undesired component (Fig. 6b).

1.2.7. Limits on amplification

When the modulation of an amplifier frame becomes large, it activates gain-control processes prior to motion

[4] Cavanagh and Anstis (1991) use a method of millisecond advancement/retardation of specific colors to measure and null color motion differences. Such a method is not practicable with the raster displays that are in common use; the 0° sandwich display proposed here is a more practical alternative.

perception. As gain-control is intrinsically slower than the controlled signal (otherwise a signal would cancel itself), the impact is felt by subsequent frames. For maximum effectiveness, amplifier frames should not occur first in the sequence, and should not be too large. The asymmetry between small-large versus large-small modulation sequences has recently been documented by Morgan and Chubb (1999); it was already evident to a small degree in van Santen and Sperling (1984). When the amplifier modulation becomes too large, there is less amplification than for intermediate modulations (see Allik & Pulver, 1995, and Section 2.1).

1.2.8. Outline

Here we apply sandwich displays and the amplification property to elaborations of the method to produce 'pure' displays that stimulate only the first-order (van Santen & Sperling, 1984), second-order (Werkhoven et al., 1993), or third-order motion perception systems. The procedure can be used to purify motion displays of any contrast amplitude to a high degree. Several issues, such as the relative advantages of the amplifier sandwich method compared to the minimum motion method, the actual amplifications achieved, and the relation of amplification limits to internal noise and to gain control, will be considered later.

2. Applications

In this section, we apply the amplification principles in various ways to produce 'pure' stimuli that stimulate only one of the motion systems.[5]

2.1. Removing first-order contamination from second-order motion displays: evaluating and tuning the amplifier sandwich paradigm

In displays that are intended to stimulate only the second-order motion system, there is a problem if first-order motion components are produced by early perceptual processes that distort the representation of luminance. Examples of such a process are a saturating nonlinearity (Scott-Samuel & Georgeson, 1999) and the greater neural representation of blacks than whites (Lu & Sperling, 1999b). The aim is to produce displays that compensate for these early visual distortions by cancelling the 'contamination' they produce.

2.1.1. Outline

The amplifier sandwich method is used to produce 'pure' stimuli as follows: (1) a real, precisely measured physical contamination is added to the carrier texture, and used in a sandwich display to calibrate and evaluate the method itself, i.e. to determine how sensitive the sandwich method is in detecting and removing the real distortion components. (2) The same method is applied to remove unknown distortion components from the candidate second-order stimuli. (3) The purified stimuli are used to test the perception of (second-order) motion. (4) To determine the amplification factor, independent of the assumption that — near threshold — $m_e m_o$ is a constant, it is efficient to add a control condition in which motion direction threshold $\hat{m}_{o=e}$ is determined when m_e is constrained to be equal to m_o. Typically, the control condition is conducted first because it determines the threshold product $m_e m_o$ and thereby the range of amplifier modulations m_e that will be most effective in steps 1–3.

Section 2.1 is concerned only with the evaluation of the amplifier sandwich method, i.e. the demonstration of its effectiveness. Section 2.2 describes its actual application.

2.1.2. Stimulus sequence

The 'second-order motion stimulus' ultimately to be used to stimulate the second-order motion system is shown in Fig. 1a and b. It is a random noise carrier with a squarewave contrast modulator that translates 90° between frames. Contrast-modulated noise is the stimulus of ultimate interest, However, to evaluate the sensitivity of the amplifier sandwich method in detecting luminance modulations, it is much more convenient to use a carrier without contrast modulation (Fig. 1p and q) that has the same mean amplitude as the contrast modulated noise.[6] A carrier without contrast modulation (Fig. 1p and q) has the advantage that, when there is no added luminance signal, motion direction is completely ambiguous.

2.1.3. Stimuli

The aim is to choose a value of m_e and to determine a psychometric function as m_o is varied for the stimuli of Fig. 1p and q. Specifically, stimuli whose motion-direction is to be detected are squarewave luminance modulations with amplitude m_o in odd frames and m_e in amplifier even frames (the same basic concept as van Santen & Sperling, 1984). Consecutive frames translate by 90° in a consistent direction. The starting phase is

[5] First reported at ARVO, 1999 (Lu & Sperling, 1999a).

[6] Previous experiments (Lu & Sperling, 1996a,b) considered masking of motion stimuli by contrast-modulated visual noise. Imposing a contrast modulation (with spatial frequency f_m) on noise had no effect whatsoever on its ability to mask a (second-order) motion stimulus of the same spatial frequency f_m. To study the effect of contrast-modulated noise on masking first-order motion, one must first remove (or greatly reduce) the first-order contamination in the visual noise so that it is direction-neutral in a sandwich display as in Fig. 1c,d. This is an considerable complication and an unnecessary refinement for the present purpose.

Table 1
Amplification of luminance motion in dynamic random noise: threshold percent modulations m_e and m_o for paired even (amplifier) and odd (test) frames for different values of m_e and the resulting amplification factors

m_e	m_o	Amplif	$\sqrt{m_e m_o}$
Subject: GA			
4.5	4.5	1	4.5
4.0	3.8	1.2	3.9
8.0	1.5	3.0	3.5
16.0	0.75	6.0	3.5
32.0	1.25	3.6	6.3
Subject: ZL			
3.0	3.0	1	3.0
4.0	3.0	1.0	3.5
8.0	0.83	3.6	2.6
16.0	0.50	6.0	2.8
32.0	0.90	3.33	5.3

random. A new sample of a random noise texture is added to each frame (dynamic random noise).

For each value of $m_e = m_o$ one of four stimuli is presented: (1) the stimulus shown in Fig. 1p and q or (2) its mirror image (i.e. motion in the opposite direction) or (3) the stimulus of Fig. 1p and q with a 180° phase shift in the modulation m_e (i.e. motion in the opposite direction) or (4) the mirror of (3). The odd frame modulation m_o is varied over the full range of values from negative (180° phase shift) to positive.

Stimuli were presented using a monochrome computer monitor with an interface that yielded 6144 levels of intensity (12.6 bits) with a mean luminance level of 40 cd/m^2. The grating size was $5.86 \times 2.93°$ (256×128 pixels). Individual pixels were $(0.046°)^2$ (2×2 pixels); the carrier noise was binary with contrasts of $\pm 50\%$. The period of the squarewave luminance grating was 1.47°, the temporal frequency was 7.5 Hz. The refresh frequency was 120 Hz, and individual frames translated 90° after every four refreshes.

On a trial, one of typically ten equally spaced values of m_o is chosen at random for presentation, and then one of the four stimulus types. (In this particular experiment, there is perfect symmetry between positive and negative values of m_o, but this is not always the case.) Thus, there are 40 possible stimulus types. Observers judge the direction of motion. The outcome of the procedure is a psychometric function of the type shown in Fig. 2a. The probability P_c of motion-direction judgments that are consistent with a positive value of m_o as a function of m_o goes from zero to one as m_o goes from large negative to large positive values. Necessarily, the expected value of P_c is 50% for $m_o = 0$. The statistic of interest is the difference between $m_{o,75}$, the value for which $P_c = 75\%$ and $m_{o,25}$, the value for which $P_c = 25\%$, which is 2 JNDs.

Once a psychometric function has been determined for a particular m_e, the procedure can be repeated for a new choice of m_e. The actual design, however, was a complete mixed list: all possible stimuli for all possible combinations of m_o and m_e were tested jointly, with a completely random selection on each trial.

2.1.4. Control experiment (Fig. 1p and q)

The procedure is identical to that described above except that instead of being fixed, m_e covaries with m_o: either $m_e = m_o$ or $m_e = -m_o$. This is an ordinary threshold determination for a translating luminance squarewave in the presence of a masking texture. It yields the value of $m_{o=e,75}$.

2.1.5. Results

Because of symmetry, $-m_{o,25}$ and $m_{o,75}$ are estimates of the same quantity, the modulation threshold, which is designated simply as $\hat{m}_o$, and estimated by the average of $-m_{o,25}$ and $m_{o,75}$. Table 1 shows the estimated threshold m_o for the various values of m_e, the first row being the motion threshold $\hat{m}_{o=e}$ obtained in the control experiment in which $\hat{m}_o$ was constrained to be equal to m_e. Column three, amplification, presents the threshold $\hat{m}_{o(e)}$ for a particular $\hat{m}_o$, m_e combination divided by $\hat{m}_{o=e}$ in the control condition (first row). Table 1 shows that a maximum amplification of 6 is reached by each observer when the amplifier frames m_e were $4–5 \times$ threshold $\hat{m}_{o=e}$.

Column 4 of the table presents the square routes of the threshold $m_o m_e$ products; $m_o m_e$ has the dimensions of power, motion power. The harmonic mean threshold modulation $\sqrt{m_o m_e}$ is shown rather than power $m_o m_e$ because it is more easily understood.

If the multiplicative property held perfectly, all the numbers in Table 1, column 4, would be identical. These products $m_o m_e$ can be regarded as inverse indexes of the stimulation efficiency of $m_o m_e$ in exciting the visual system. For both observers, $m_o m_e$ pairs with amplifier modulations that are two or three times larger than the threshold modulation $\hat{m}_{o=e}$ yield the highest stimulation efficiency. Still greater amplifier modulations m_e (not shown here) result in a severe loss of efficiency. Very large m_e reduce the amplification to much less than 1, that is, they mask rather than amplify the test frames.

The loss of sensitivity with very large amplitude amplifier frames is due largely to gain control mechanisms prior to the motion computation. The small reduction of $m_o m_e$ (an increase of sensitivity) in the range of maximum amplification is more difficult to explain. One possibility is that with larger inputs, perceptual filters become more finely turned. In that case, the large modulation in the amplifier frames would serve not only to amplify odd frames but also to sharpen perceptual filters, and thereby reduce the amount of stimulus noise that is unavoidably amplified with unequal versus equal m_o, m_e.

2.1.6. Tuning the amplifier sandwich display for static noise

According to the Reichardt model, static stimuli of any kind can be added to a motion stimulus without having any effect on motion thresholds. Because motion displays typically turn on and off in relatively brief time periods, there is no such thing as a truly static stimulus. However, for sampled motion displays with 90° phase shifts between frames, and $4n+1$ frames per display (n = positive integer), it has been shown (Sperling & Lu, 1998b) that even the most damaging of static displays, turned on and off in synchrony with the motion stimulus, can be virtually invisible to a Reichardt detector. Here we consider a static noise carrier that has a modulation amplitude of $\pm 50\%$, which is almost 500 times the amplitude of the threshold of a first-order motion stimulus against a neutral gray background. As a masking stimulus, the static noise raises the threshold modulation of luminance motion about $5\times$ as compared to a neutral gray background.

Using procedures and stimuli identical to those described above, except that the noise carrier texture remained unchanged throughout a trial, thresholds were measured for various m_o, m_e combinations (Table 2). The threshold amplitude is $4\times$ lower with the static than with dynamic noise. The maximum amplification factor is 7.5–8 for the two observers. These factors are reached with amplifiers somewhat less than $7.5–8\times$ threshold, again indicating a slight gain in efficiency before the drop in efficiency at about $10\times$ threshold.

Table 2
Amplification of luminance motion in static random noise: threshold percent modulations m_e and m_o for paired even (amplifier) and odd (test) frames for different values of m_e and the resulting amplification factors

m_e	m_o	Amplif	$\sqrt{m_e m_o}$
Subject: HK			
1.20	1.20	1.0	1.20
1.0	0.80	1.5	0.9
2.0	0.40	3.0	0.9
4.0	0.15	8.0	0.8
8.0	0.22	5.5	1.3
16.0	0.75	1.6	3.5
32.0	1.0	1.2	5.6
Subject: ZL			
0.75	0.75	1.0	0.7
1.0	0.50	1.5	0.7
2.0	0.16	4.7	0.6
4.0	0.10	7.5	0.6
8.0	0.25	3.0	1.4
16.0	0.25	3.0	2.0
32.0	0.72	1.0	4.8

2.1.7. Psychometric functions and the precision of estimating motion components

Fig. 2a shows an idealized psychometric function of the sort obtained in these experiments. It is obvious that the greater the amplification (the smaller the threshold), the steeper is the psychometric function. The increase in steepness of the psychometric function is a direct consequence of the definition of the psychometric function and of the threshold.[7] The importance of an increase in steepness of the psychometric function is that the precision of estimation of the point where it crosses the line $P_c = 50\%$, the median of the underlying probability density function, is directly proportional to the slope of the psychometric function (see Section 3.1). And the precision of estimating m_c, the crossing point, is the critical issue in precisely cancelling possible contamination motion components.

2.1.8. Conclusion

Considerable amplification of small luminance signals in dynamic and static noise texture backgrounds can be achieved by use of a sandwich method. The greatest amplification occurs with amplifier modulations that are about four to six times the normal motion threshold (in that environment).

2.2. Removing first-order contamination from second-order motion displays: creating 'pure' second-order stimuli

The aim is to create a moving contrast-modulation (second-order stimulus, e.g. Fig. 1a and b) that contains no motion components to activate the first order (luminance) motion system, even after luminance distortion which occurs early in visual processing. There are some suggestions that second-order motion would be weak or absent in such stimuli (e.g. Smith & Ledgeway, 1997; Taub, Victor, & Conte, 1997). Both static and dynamic stimuli are tested as described below. The amplifier sandwich method is used to remove the undesired first-order components. The contamination-detection and nulling procedures were implemented on a Proxima DLP-4200 projector controlled by a Macintosh PowerPC computer running Video Tool Box graphics software (Pelli & Zhang, 1991). The Proxima DLP-4200 projector was set into its achromatic mode with a refresh rate of 67 frames/s. The gamma of the projector was re-set to 1.0 using a proprietary program. The voltage-luminance response of the DLP-4200 is linear

[7] Psychometric functions often are plotted against the logarithm of stimulus intensity. In this case, the center of the psychometric function is taken as the point of zero intensity. On a logarithmic intensity scale, an increase in sensitivity (change in slope on linear scale) would be manifest as a translation to the left, often with an invariant shape of the psychometric function.

within measurement error. In the chromatic mode, each RGB channel of the projector produced 256 (8-bit) color levels. In the achromatic mode, the projector produced 768 evenly spaced gray levels. The observer viewed the display (0.6×0.4 m^2) on a reflecting screen at a distance of 3.0 m. The dynamic range of the projection system was 10.3–151.6 cd/m^2; the background luminance was 81.0 cd/m^2.

The odd frames (test frames to be calibrated) of the displays consisted of a carrier (binary random noise, pixel size $= 8.4 \times 8.4$ min) upon which a square-wave contrast modulation (0.45 cd) was imposed (Fig. 1c,d). The intensity of pixel x is defined in terms of its point contrast (in percent): $c(x) = 100\%\ l(x)/l_0$, where l is the pixel luminance, and l_0 is the background luminance. In all cases, l_0 was also the mean luminance of the entire display. By choosing l_0 as the midpoint of the range of displayable luminances, the theoretical contrast range of the monitor is $[-100\%, +100\%]$. In high contrast stimulus regions, pixel contrast was -75% or $+75\%$ with equal probability; in low contrast regions, pixel contrast was -25% or $+25\%$ with equal probability. The even frames (amplifier frames) of the displays contained the same carrier (binary random pixel noise) but instead of contrast modulations they contained luminance modulations of the same spatial frequency. In the high luminance regions, pixels took contrasts of $+75\%$ or -25% with equal probability; in the low luminance regions, pixels took contrasts of $+25\%$ or -75% with equal probability. This produces a luminance modulation (in terms of point contrast) of $m_e = \pm 25\%$ in the amplifier frames. This value of m_e is slightly too big to yield the maximum amplification, but it had the virtue of requiring only four luminance levels for a stimulus, which ultimately simplified the calibration problem. The phase shift between successive frames was 90°. The display was shown at 7.5 Hz.

If there were no luminance contamination, no motion system would be activated by the display of Fig. 1c and d. Within odd frames alone and within the even frames alone, the phase shift is 180°, which is motion-ambiguous. The odd frames address only the second-order system, the even frames, only the first-order system. However, Lu and Sperling (1999b) found that, in all the texture displays they tested, black pixels had greater perceptual effectiveness than equivalent white pixels. This means there is an early nonlinearity in the visual system that diminishes positive contrasts relative to negative ones. This kind of luminance distortion would produce a first-order contamination in the texture frames. After early visual processing, all frames would contain a luminance stimulus, and first-order motion would be easily perceived.

Two versions of the display were tested. In the static carrier version, the random binary noise pixel pattern was the same across all five frames of the display. In the dynamic carrier version, a new random binary noise pixel pattern was chosen on every frame. With a static carrier, only the modulator changes from frame-to-frame; so all changes in pixel intensity are relevant for motion. With a dynamic carrier, most of the changes in pixel intensity are due to changes in the carrier (which are irrelevant to motion) and not in the modulator.

Both static- and dynamic-carrier versions showed strong apparent motion. The apparent motion was nulled by adding a luminance modulation m_o to the second-order grating (white aligned with high contrast region to cancel the black advantage). A method of constant stimuli was used to estimate the modulation amplitude that would render the modified test grating completely motion-ambiguous when alternated with the amplifier frames (e.g. Fig. 2a). Four observers were tested with both gratings.

2.2.1. Results

For the four observers, the cancelling luminance-modulation varied from 3.0 to 4.5%. There was no statistically significant difference in the nulling modulation required for static and dynamic carriers (contrary to the conjecture of Smith & Ledgeway, 1997). In second-order stimuli, with all first-order contamination removed, the direction of motion remained completely obvious.

2.2.2. Estimating amplification

Amplification is estimated by comparing the slopes of psychometric functions, e.g. data of the form of Fig. 2a: P_c, the probability of a direction-consistent motion judgment versus m_o, the odd-frame modulation amplitude, while even-frame modulation amplitude m_e is fixed. In the demonstration measurements of amplification in Section 2.1, the psychometric functions were constrained to go through the point (0, 50%) because the odd-frame texture did not contain any intrinsic luminance variations that could support first-order motion. Here, however, the odd frame texture has a contrast modulation that, after the compressive nonlinearities of retinal processing, contains intensity modulations that subsequent visual processes cannot discriminate from an equivalent first-order 'contamination'. In consequence, except for statistical variability, all the psychometric functions pass through a point (m_c, 50), where m_c represents the modulation amplitude of the first-order contamination (Fig. 2a).

The slope of a psychometric function for a fixed m_e is determined by the distance between $m_{o,25}$ and $m_{o,75}$, i.e. the values that yield the 25% and 75% points, $P_c(m_{o,25}) = 25\%$ and $P_c(m_{o,75}) = 75\%$. For a given m_e, the threshold m_o^{jnd} is defined as $m_o^{jnd} = (m_{o,75} - m_{o,25})/2$ (see Fig. 2a). The slope of the psychometric function is $dP_c(m_o)/dm_o \approx 25\%/m_o^{jnd}$.

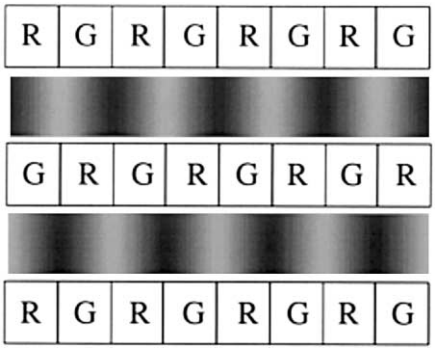

Fig. 7. Procedures used to detect first-order contamination in nominally isoluminant red–green gratings. Odd frames, the red–green gratings. Even frames, the luminance amplifier (Anstis & Cavanagh, 1983).

Amplification for a particular m_e is the ratio of slopes of two psychometric functions. The numerator is the slope of the psychometric function derived with me; the denominator is the slope of the reference psychometric function, derived with $m_e = \hat{m}_{o=e}$ [A complication in the present experiment (Section 2.2 versus the previous Section 2.1) is that contrast-modulated noise contains a luminance component; this luminance contamination must be removed before the threshold determination of $\hat{m}_{o=e}$ (see Section 2.1) in which ($|m_e|$ is constrained to be equal to ($|m_o|$.] From the definition of slope, it is evident that amplification is the ratio of the corresponding two thresholds: $m_{o,ref}/m_o$. The slope ratio indicates that the actual amplification in this experiment was about 5–6.

2.2.3. Discussion

Similar experiments and results in the removal of first-order contamination from second-order motion stimuli have recently been reported by Scott-Samuel and Georgeson (1999). In addition to studying second-order motion produced by moving modulations imposed on noise carriers, they also used high spatial-frequency sinewave carriers. The disadvantage of a sinewave carrier, noted by the authors, is that the modulator produces real physical first-order movement components that interfere with the measurement of the visually produced first-order motion components. The random noise carriers and modulators used in these experiments have been proved not to contain systematic first-order components (Chubb & Sperling, 1989a,b). Any first-order motion that can be cancelled must have been visually produced.

Currently, purified stimuli are being used to study transducer properties of second-order system (Chubb, Lu, & Sperling, 1999), motion summation, cross-adaptation, and other phenomena that require pure second-order stimuli for their elucidation.

2.3. Removing luminance contamination from isoluminant chromatic motion displays

In studying chromatic motion perception, it is critical that the chromatic stimuli be truly isoluminant with respect to human motion perception. Anstis and Cavanagh (1983) first used an amplifier sandwich display and nulling procedure to remove such luminance contamination. However, they were not aware of the amplifier property of sandwich displays. In the amplifier version of the procedure, we create calibration displays consisting of five consecutive frames. The odd frames are test frames composed of the nominally isoluminant red–green sinewave gratings to be calibrated; the even frames are amplifier frames composed of sinewave luminance modulations. The phase shift between successive frames is 90° (Fig. 7).

No consistent motion would be perceived in a color-luminance sandwich display (Fig. 7) if the 'isoluminant' red–green frames were truly isoluminant. If there were residual luminance contamination in the 'isoluminant' frames so that, for example, the green areas were slightly less luminous than the red areas, motion would be seen in a consistent leftward direction; if green were more luminous than red, motion would be seen in the rightward direction, the directions reflecting the relative phases of the luminance components which are now present in all five frames.

Measurements of the effective amplification in the sandwich procedure for producing red–green isoluminant motion produced different results than in experiment 1. The maximum amplification achieved was about 2.0, and this occurred at low contrasts, about 4–8%, of the luminance amplifier frames. The procedures for measuring the effectiveness of amplification (e.g. Tables 1 and 2) consist of the first-order sinewave amplification sandwich procedure (Fig. 3) plus different added masking stimuli–noisy textures in experiment 1, red–green gratings in isoluminance calibration. That color masking appears to follow somewhat different principles is unexpected and under investigation.

The isoluminance calibration procedure was implemented on a Macintosh computer. All displays were shown on an Apple 1710 multisync color monitor controlled by a 30-bit Radius Thunder 1600/30 graphics card in a 7500/100 PowerPC Macintosh running Psych Toolbox (Brainard & Pelli, 1998). This apparatus was different from that used in the previous experiment because it was necessary to have ten bits per channel to produce truly isoluminant chromatic stimuli. The monitor, with a refresh rate of 60 frames/s, was calibrated to produce DKL coordinates following standard procedures (Brainard, 1986) using a Tektronix J17 photometer with a J1820 chromaticity head. The background was set at 17.4 cd/m^2 at chromaticity $(x, y) = (0.294, 0.314)$ for the CIE Standard Observer.

The chromatic 'red–green' test grating is a 0.5 cpd sinewave modulation along the $L-M$ axis in the DKL space (Derrington, Krauskopf, & Lennie, 1984). The luminance amplifier grating is a sinewave modulation of the same spatial frequency along the $L+M$ axis in the DKL space. To find a 75% threshold for a purified test grating, an initial (contaminated) test grating was set at an amplitude of about 1.5 threshold (i.e. about 90% detectable) and then calibrated by means of the motion amplification procedure to remove the residual luminance component. (In our experience, even after the general isoluminant calibration has been performed, the motion-amplification procedure discovers that about 1/3 of the nominal isoluminant modulation at threshold actually contains residual luminance contamination. Moreover, the final calibration is slightly different at every spatial frequency, temporal frequency, and contrast amplitude, as discussed by Anstis & Cavanagh, 1983.) Extensive isoluminance calibration was carried out for four observers and at four different temporal frequencies (0.92, 1.84, 3.75 and 7.5 Hz) and several contrast amplitudes near threshold. The amplitudes of the resulting isoluminant threshold modulations along the $L-M$ axis for the four observers were 1.2% (−0.5%), 1.35% (−0.5%), 2.4% (−0.5%), and 1.5% (−1.0%) at 7.5 Hz and the average thresholds for the four different temporal frequencies were 0.44% (−0.6%), 0.48% (−0.13%), 0.68% (−0.18%) and 1.6% (−0.63%), with the numbers in parenthesis indicating the amount of luminance correction in each condition.

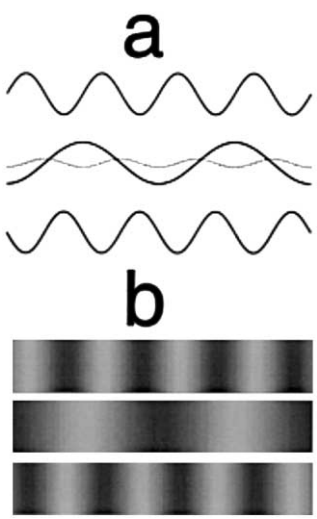

Fig. 8. Sandwich procedure used to detect $2f$ distortion in sinewave motion displays. To detect second harmonics in a nominally pure sinewave grating, a $1f$ sinewave grating is sandwiched between two $2f$ sinewave gratings. The phase shift between the two $2f$ gratings is 180°. If there were second harmonics in the internal representation of the $2f$ grating (as indicated by the dotted curve in (e), apparent motion in a consistent direction is perceived. To remove $2f$ components from the internal representation of $1f$, various amounts (including 0) of $2f$ components are added to the $1f$ grating until no consistent motion is perceived. The procedure can be iterated to remove $4f$ and higher harmonics.

The properties of these truly isoluminant 'red–green' stimuli are described elsewhere (Lu, Lesmes, & Sperling, 1999) and are briefly summarized here as observed in each of four observers. (1) By appropriate adjustment of chromaticity, the appearance of chromatic motion could be adjusted from smooth, continuous motion to complete motion standstill (a quickly moving grating appears to be standing still). (2) The temporal frequency tuning function (threshold amplitude versus frequency) exactly matched the previously measured tuning functions for third-order motion stimuli. (3) Pure chromatic stimuli failed the pedestal test (Lu & Sperling, 1995a). (4) Monocular stimuli and interocular stimuli (in which each eye individually receives only ambiguous motion) had the same thresholds within measurement error. Each of these properties, and all together, indicate that the movement of isoluminant red–green gratings is perceived by the third-order motion system.

2.4. Measuring distortion components in first-order sinusoidal gratings

Lu and Sperling (1999b) report that, in ten representative types of first-order stimuli, decremental black areas have approximately 25% greater perceptual magnitude than incremental white areas that have the same absolute physical deviation from the background.[8] The perceptual distortion of sinewaves caused by black–white asymmetry is equivalent, in a Fourier series representation, to adding even harmonics to the original sinewave. That is, when the original sinewave frequency is f, the black–white asymmetry produces a waveform that, in addition to f, contains frequencies $0f$, $2f$, $4f$, $6f$,... . The magnitude of these distortion-produced sinewaves is predicted from the black–white asymmetry.

A sensitive test for the presence of the $2f$ distortion product is the three-frame motion sequence shown in Fig. 8. Frames 1 and 3 are the amplifier frames: Frame 1 is a $2f$ spatial sinewave; frame 3 is its negative. These frames alone are motion-ambiguous. Frame 2, the test frame, contains a $1f$ sinewave with a 45° phase shift relative to frame 1. When the if sinewave contains a $2f$ distortion product, the $2f$ phase shift is 90°, and motion will be seen in the rightward direction if black has a bigger representation than white, and in the leftward direction otherwise. When there is no distortion

[8] Early visual nonlinearities in which the magnitude of visual responses to decrements is larger than the response to equal-magnitude increments are reported for: turtle cones (Baylor, Hodgkin, & Lamb, 1974), human cones (He & MacLeod, 1998), brief threshold flashes (Boynton, Ikeda, & Stiles, 1964, and many subsequent authors), and ten other perceptual responses (Lu & Sperling, 1999b).

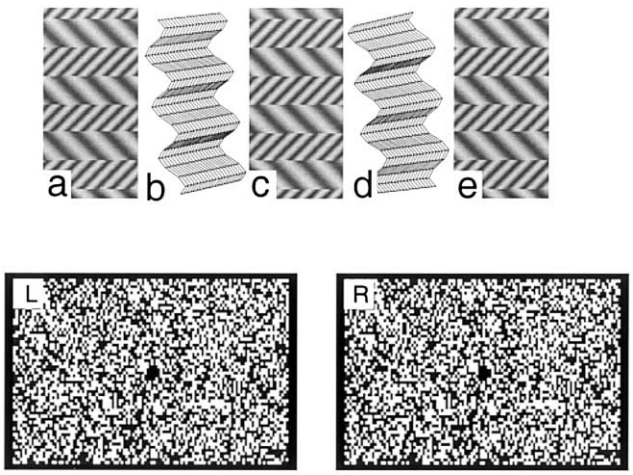

Fig. 9. Removing bias in third-order motion: the depth-texture sandwich display. The top row shows a sequence of five consecutive frames (a–e); each is displaced by 90° from the previous one. Overall frame size is $7.43 \times 5.94°$. The depth stereograms (a, c, e) are indicated schematically; an actual stereogram is shown below (L, R). The most salient features of the depth frames are the near peaks. To balance the salience of the coarse and the fine stripes in equal attention conditions, the contrast of the coarse stripes is varied while the contrast of the fine stripes is fixed. When the coarse stripes are more salient, the direction of apparent motion follows the space-time trajectory from coarse stripes to near peaks of the depth gratings. The opposite direction of motion is perceived when the fine stripes are more salient. When no consistent motion direction can be reported, the coarse and the fine stripes are equally salient. (L, R) Left and right eye images of a stereogram illustrating the bottom half of one frame of the depth stimulus. Viewing the left panel with the left eye and the right panel with the right eye shows 1.7 (of 3.7) cycles of a horizontally oriented grating (Lu & Sperling, 1995b).

product, there is no common frequency between consecutive frames, and therefore there is no consistent apparent motion.

Three observers viewed the test sequence of Fig. 8 using the DLP-4200 projector described in experiment 1. The spatial frequency of the fundamental sinewave was $f = 0.60$ cpd; the spatial frequency of the second harmonics was $2f = 1.2$ cpd. Each frame of the motion sequence lasted 50 ms. The contrast of all the gratings was 4%, which is quite small and, according to most theorists, should be relatively free of intensity-compression distortion.[9]

To compensate for the presumed early visual black–white distortion in the representation of luminance, an opposite distortion ('black attenuation') was produced in the stimuli. Specifically, let the luminance of a pixel be l, and the mean luminance be l_0 and let $\Delta l = l - l_0$. The point contrast of 'black' (negative) pixels (pixels with $\Delta l < 0$) was attenuated by α, the black-attentuation factor. That is, the actual displayed luminances l' of pixels with $\Delta l < 0$ was $l' = l_0 + \alpha \Delta l$; white (positive) pixels were not transformed, $l' = l_0 + \Delta l$.

We corrected for stimulus distortion by black attenuation rather than by simply adding second harmonic because black attenuation is a reasonable antidote to the actual visual attenuation process and because it had already been used by us in many similar psychophysical tests (Lu & Sperling, 1999b). In this particular instance, black attenuation is, for practical purposes, equivalent to adding second harmonic because the higher harmonics (4th, 6th…) produced by black attenuation are insignificant in such low-contrast stimuli.

The black-attenuation factor α was varied to determine the point of minimum motion, i.e. where the observer's internal representation of black in this particular stimulus was of precisely the same magnitude as that of white. Even harmonics higher than $2f$ were not tested in this paradigm because they would have phase shifts of 180 or 360°. For three observers, the average point of minimum motion occurred when black was attenuated by $\alpha = 0.72$. Black attenuation by a factor α reduces the amplitude of the fundamental by $(1 + \alpha)/2 = 86\%$, and it produces a $2f$ component of amplitude $2(1 - \alpha)/(3\pi) = 5.9\%$. The amplitude of $2f$ relative to $1f$ is $0.059/0.86 = 0.069$. This 7% distortion of a sinewave grating that has only 4% contrast modulation implies an actual $2f$ modulation amplitude of 0.28%. The large black–white distortion ($\alpha = 0.72$) generates only a small $2f$ component (modulation amplitude 0.28%); nevertheless, the $2f$ component is easily measured and removed.

In principle, one can reiterate the above procedure to remove $4f$, $6f$ and even higher harmonics from the sinewave to produce a $1f$ sinewave grating that is perceptually pure (devoid of harmonics) for the first-order motion system. Such purification might become necessary when displays utilize higher-amplitude moving sinewave stimuli or frame-to-frame displacements other than 90° because of extreme sensitivity of the first-order motion system to low-contrast motion.

Ordinary sinewave gratings appear to have dark stripes that are quite obviously wider than light stripes. However, a perceptually pure sine stimulus for the motion system is not necessarily a pure sine for the shape–texture system. That would require a spatial calibration (see Section 2.7, below). For spatially purified sinewave gratings, the apparent widths of black and white stripes were found to be equal within measurement error (Lu & Sperling, 1999b). Insofar as the black-attenuation factors measured for grating appearance and for motion are similar, it suggests a common origin for black-white asymmetry.

[9] Near threshold, detection typically follows a square law as a function of stimulus intensity. However, the square law occurs at a higher level of processing than the representation of the inputs to the motion computation. Predictions of how two stimulus components combine in determining motion-direction judgments indicate that, when stimulus contrast is small enough to elude contrast gain control, stimuli are represented very accurately, i.e. linearly (van Santen & Sperling, 1984; Lu & Sperling, 1995a; Lu & Sperling, 1996b).

2.5. Removing second-order contamination from third-order motion displays

Lu and Sperling (1995b) created a novel sandwich display to stimulate only the third-order motion system (Fig. 9). Odd frames are texture gratings, even frames are stereo depth gratings. Perceiving motion in these displays requires combining the salience information (e.g. figure vs. ground) from the two different stimulus types. Additionally, Lu and Sperling (1995b) intended to produce a motion display in which the motion would be invisible to the first- and second-order systems, and would be visible to the third-order motion system only when selective attention was directed to a particular feature.

Three calibrations are desirable in third-order displays used to assay selective attention: first-order and second-order motion should be eliminated. Neither of these is absolutely essential, but the sensitivity of the method is enhanced when extraneous motion signals are eliminated. Finally, in equal-attention conditions, no consistent (third-order) motion should be reportable by the observers.

Sandwich displays that have been used to study selective attention to black (or to white) spots are illustrated in Fig. 5. Fig. 5a illustrates a display in which even frames, the amplifier frames, are composed of a high-amplitude contrast-modulated texture grating (i.e. with alternating bands of high and low texture-contrast). Odd frames, the test frames, consist of alternating bands of white 'Mexican hat' spots and black 'Mexican hat' spots. These spots are composed of a center and surround of opposite contrast so that, for each kind of spot, the integrated luminance over the entire spot equals the background level. Before beginning the procedure, the (first-order) luminance modulation is eliminated in both types of frames by means of a first-order calibration.

For eliminating second-order motion components, the amplifier frames have a second-order modulation of about $10\times$ the threshold for texture modulation. Insofar as there is any second-order modulation in test frames (e.g. if the black spots were blacker than the white spots were white) it would lead to apparent motion in a consistent direction. Because of the greater effectiveness of black than white, minimum motion typically requires reducing the amplitude of the black spots to about 90% of the amplitude of the white spots, somewhat different amounts depending on the observer (Solomon & Sperling, 1994). (Changing the amplitude of the spot involves the amplitude of both the spot and its immediate 'Mexican hat' surround in order to leave the first-order calibration unperturbed.) Once the test display is black–white balanced and therefore motion-ambiguous, it no longer contains a second-order motion component.

A complication in this procedure is that, although the calibration is conducted under conditions of neutral attention, the observer may be selectively attending more to one color (black/white) of the spot than the other. Selective attention to a color makes that color more effective for third-order motion (Lu & Sperling, 1995b; Blaser et al., 1999) so the balance would be achieved between a combination of second- and third-order motion. To achieve a perfect cancellation of second-order motion without any contribution of third-order motion, it would be necessary to conduct the calibration at a sufficiently high frequency, e.g. 12 Hz, so that the contribution of third-order motion is negligible. (The typical corner frequencies of second and third-order motion are, respectively, 12 and 4 Hz.)

Solomon and Sperling (1994), in a different context, used a 'minimum motion' procedure to remove fullwave (second-order) contaminations from the same black–white Mexican hat displays. Illustrated in Fig. 5c,d, their procedure varied the ratio of the contrast of the black dots versus that of the white dots to search for a particular black/white ratio at which minimum amount of motion is perceived. They took that particular black/white ratio to indicate successful removal of fullwave (second-order) contamination in the black- and white-Mexican-hat frames. While their procedure enabled them to remove much, perhaps most, second-order contamination from their displays, it has two problems that are resolved by amplification sandwich displays. (1) The Solomon–Sperling procedure is less sensitive because second-order contamination below its own threshold fails to cause motion, whereas in the sandwich procedure a second-order contamination of 1/8 of threshold (Table 2) causes detectable motion. (2) In order to make it possible to observe changes in motion strength in their displays, Solomon and Sperling (1994) had to keep the stimulus contrast low enough that the direction-of-motion judgments were less than about 90% correct. Then, they had to linearly extrapolate their calibration results to the high contrast stimuli actually used in their experiments. In contrast, the amplification sandwich procedure can be used to directly calibrate motion displays of any amplitude to a certifiable purity.

The principal result with stimuli that are motion-balanced under neutral attention conditions is that selective attention to a feature (either the black spots or the white spots in the present example) can completely determine the direction of perceived motion (Lu & Sperling, 1995b) — the same stimulus appears to move in one direction with attention to black spots and in the opposite direction with attention to white spots. In the case of selective attention to color, the increase in effectiveness of a color (the attentional amplification) produced by selectively attending to red or to green has been measured quantitatively and found to be 25% (or

more, depending on the observer and color). The increase in color effectiveness is only in terms of the color's salience contribution to third-order motion; color appearance itself is not changed by selective attention (Prinzmetal, Amiri, Allen, & Edwards, 1998; Blaser et al., 1999).

2.6. Removing third-order bias from third-order motion displays

This example provides perhaps the clearest case of third-order motion in a sandwich display (Fig. 9). The even, amplifier frames are dynamic random-dot stereograms. Each frame individually is composed of random-intensity pixels. The correlation between the pixels in the left and right eyes is arranged so that together the pixels describe a horizontal, corrugated grating in depth. The depth modulation is far above threshold, hence these frames function as third-order amplifiers, with the near areas perceived as figure (high salience) and the distant areas as ground (low salience). The odd, test frames are alternating patches of high frequency and low frequency sinewave gratings. Because there is nothing common between the random-dot stereo frames and the texture patches, there is no possibility of perceiving first or second-order motion. However, observers will perceive motion in a direction that depends on the relative contrast amplitude of the high and low spatial-frequency gratings. Under neutral attention instructions, adjusting the relative amplitude of the high and low spatial-frequency sinewave gratings to null apparent motion determines the high frequency/low frequency amplitude ratio for third-order motion neutrality. (Under neutral attention, the third-order system is low-pass in spatial frequency; the higher the patch spatial frequency, the higher the contrast must be for a motion null.)

The procedure was carried out on an Ikegame monitor driven by an IBM PC compatible running Runtime Library. The overall size of the stimulus as viewed by the observers was $5.94 \times 2.97°$. The contrast of the fine texture gratings was 40%. At equal salience, the contrast of the coarse gratings was 26%. The coarse stripes were sinewaves, 2.5 cpd, the fine stripes, 5.0 cpd.

Under instructions to selectively attend to high or to low spatial frequencies, the stereo-alternating-with-texture-patch display enables the calibration of the effectiveness of attention instructions. In this case, attention to high causes reliable motion perception in one direction, attention to low spatial frequencies causes motion perception in the opposite direction. In a similar ambiguous motion paradigm, Blaser et al. (1999) used a texture-contrast modulation as the third-order motion amplifier and an isoluminant red–green grating as the test display. The calibration consisted of adjusting the saturation of red or green stripes to null apparent motion. For the third-order motion system, attention to a color was found to be equivalent to increasing its saturation by at least 25%.

2.7. Amplifying texture gratings

To what extent does the amplification principle observed in the motion system apply to texture gratings? In a formal sense, the problem of determining the direction of motion of a one-dimensional grating is exactly equivalent to the problem of determining the slant of a texture grating. This is demonstrated by inspection of Fig. 1a, which represents a space–time (x, t) plot of a one-dimensional grating modulation (Fig. 1b) that moves in 90° steps. The coordinates of the plot can be interpreted as x, t representing successive frames (as in all the previous examples) or as directly as x, y: The plots themselves (e.g. Fig. 1a,c) are x, y gratings. Can the amplification principles, applied in such instances to alternate slices of the gratings' stripes, be used to make the slant of the gratings more visible? And, if so, what does that tell us about texture perception?

2.7.1. Procedure

To investigate amplification in texture gratings, in a setting as similar as possible to the motion stimuli, the luminance square-wave modulations illustrated in Fig. 10 are used. (These experiments were conducted in collaboration with Greg Appelbaum.) The stimuli were composed of $4n$ slices (rows) of squarewave luminance modulations with amplitude m_o in odd (test) slices and m_e in even (amplifier) slices. Consecutive slices were translated in a consistent direction by 90° (Fig. 10). The task of the subject is to discriminate between the rightward-slanting stimuli of Fig. 10 and their mirror images.

Initially, the method of constant stimuli was used to determine the contrast amplitude which produced the probability of correct texture slant judgments $P_c = 75\%$ when all strips had the same contrast, i.e. when $m_o = m_e$. This is an ordinary threshold determination for texture slant discrimination.

Once the threshold modulation $\hat{m}_{o=e}$ had been determined for the squarewave that produced 75% correct slant judgments, the method of constant stimuli was repeated with various fixed amplitudes of m_e to determine the corresponding psychometric functions as m_o was varied. For each fixed value of m_e, the value $\hat{m}_o$ was determined that made texture slant judgments 75% correct.

2.7.2. Stimuli

Stimuli were presented on a monochrome computer monitor with an interface that yielded 6144 levels of intensity (12.6 bits) and a mean luminance level of 40

cpd/m^2. The grating size was $2.93 \times 2.93°$ (256×256 pixels). Individual texture elements were $(0.046°)^2$ (2×2 screen pixels). Fig. 10a,b show the gratings, approximately to scale. Two spatial frequencies of the square-wave luminance grating were tested: 1.2 and 2 cpd. The texture was shown for 50 ms at a refresh frequency of 120 Hz.

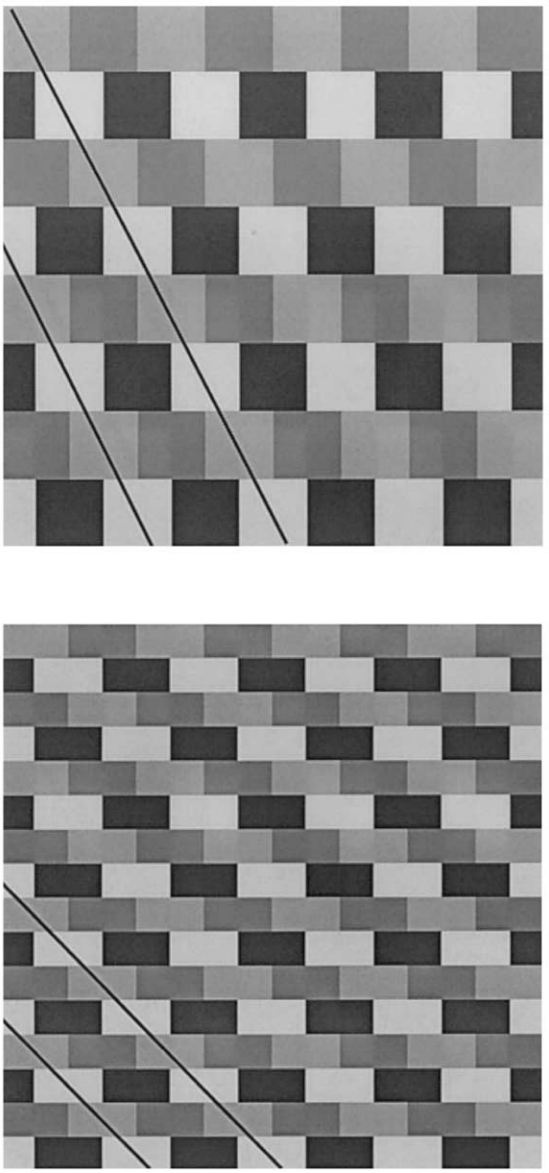

Fig. 10. Stimuli for producing texture amplification analogously to amplification in motion sandwich displays. (a) A section of right-slanting texture. Even slices are the amplifier slices with high contrast amplitude m_e; odd slices are the to-be-amplified slices with low contrast amplitude m_o. The discrimination is between the right-slanting grating and its left-slanting mirror image. (b) A right slanting grating with a slightly different geometry than (a). In the experiments, threshold contrast $m_e = m_o$ was 0.6%; the minimum (amplified) m_o for 75% correct slant discrimination was less than 0.1% (one part in a thousand).

Table 3
Slant discrimination (Fig. 10): pairs of threshold contrast modulation amplitudes in percent for amplifier slices m_e, test slices $\hat{m}_o$, the corresponding amplification factors, and the RMS orientation power

m_e	$\hat{m}_o$	Amplif	$\sqrt{m_e m_o}$
1.2 cpd, 50 ms			
0.58	0.58	1	0.58
1	0.34	1.7	0.58
2	0.088	6.6	0.42
4	0.12	4.8	0.69
8	0.70	0.23	2.36
2 cpd, 50 ms			
0.58	0.58	1	0.58
1	0.33	1.76	0.57
2	0.075	7.73	0.39
4	0.20	2.9	0.89
8	0.56	1.03	2.1

2.7.3. Results

The first line of Table 3 shows the ordinary threshold $\hat{m}_{o=e}$ for slant discrimination when $m_e = m_o$. Subsequent rows show the threshold values $\hat{m}_o$ for various m_e. Amplification is the threshold m_o (for that row's m_o, m_e combination) divided by $\hat{m}_{o=e}$. Table 3 shows that amplifications from 6 to 8 are reached by the observer when the amplifier frames m_e were about $3.5\times$ the threshold $\hat{m}_{o=e}$, and that further increases in m_e increase $\hat{m}_o$ ultimately producing masking, not amplification.

The greatest amplification, which occurs when m_e is only 3.5 times threshold $\hat{m}_{o=e}$, is a joint consequence of multiplicative amplification (by m_e) and a significant increase in efficiency — a smaller $\sqrt{m_e m_o}$. The conclusion, based on the current data, is that amplification is a potentially useful principle in exposing weak x, y texture components analogous to its use in x, t motion experiments.

2.7.4. Interpretation

We expect amplification in texture because algorithms for detecting grating orientation rely on extracting orientation energy in x, y (Knutsson & Granlund, 1983), which is computationally equivalent to extracting motion energy in x, t (Adelson & Bergen, 1985). The equivalence of motion energy models and the elaborated Reichardt motion model (which has the amplification property) was proved by Adelson and Bergen (1985) and van Santen and Sperling (1985). So, by transitivity (elaborated Reichardt model equals motion energy model equals orientation energy model), the amplification principle must hold for texture slant insofar as the slant discrimination is based on an orientation energy computation.

3. Discussion

This article illustrated the amplification principles in first-, second- and third-order motion. Based on these principles, sensitive calibration procedures were developed to reduce unwanted stimulus components to a small fraction of their threshold, i.e. to produce 'pure' stimuli in which the undesired components, even if they were an order of magnitude greater, would remain invisible.[10] As briefly summarized in each section, such pure stimuli enabled selective stimulation of each of the proposed three motion systems, best revealing the intrinsic properties of these systems.

3.1. Efficiency

How much is gained by motion amplification is illustrated in Fig. 11. Suppose the method of constant stimuli is used to collect motion-direction responses. We consider two points of the psychometric function that describes the perceived motion-direction as a function of the amount of added, cancelling luminance motion. As in the psychometric function of Fig. 2, we wish to estimate the point of maximum motion ambiguity, where the psychometric function crosses the line that describes chance performance. Suppose, in a motion procedure, data points b and b' are obtained. The range of uncertainty around each of these data points is proportional to the standard error of the mean, represented by the vertical bar. Connecting the bars from b and b' indicates how uncertainty in the proportion of responses translates into uncertainty in the location of the crossing point (the median of the underlying probability density function). The exact relation between vertical and horizontal uncertainty will be considered below; here, only the comparison between two different uncertainty intervals is considered.

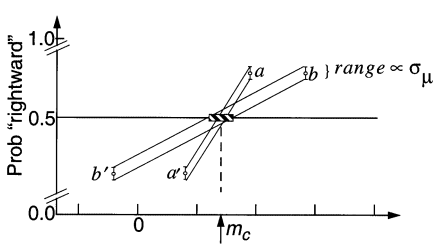

Fig. 11. The efficiency of amplification. Idealized pairs of points on two psychometric functions (Fig. 2a) obtained in sandwich procedures. Abscissa is modulation amplitude m_o of odd frames, ordinate is the probability P_r of rightward motion. $b-b'$ are baseline data from a psychometric function obtained with a small modulation amplitude m_e of even frames. $a-a'$ are obtained with $3\times$ larger m_e, i.e. $3\times$ amplified data. The bars around the data points indicate the standard error of the mean σ_μ. The range of uncertainty around the estimate of m_c diminishes in direct inverse proportion to the amplification factor α and σ_μ. To reduce σ_μ by α requires α^2 more independent observations.

In Fig. 11, data points a–a' arise when there is an amplification factor of $3\times$ with precisely the same amount of data. It is evident that the $3\times$ amplification has reduced the uncertainty interval threefold. A threefold decrease in the uncertainty interval could also be produced by a threefold reduction in the range. This would require a three fold reduction in the standard error σ_μ of the mean. Unfortunately, $\sigma_\mu = \sigma\sqrt{N}$, where σ is a constant determined by the Bernoulli distribution associated with P_c, the response probability. To decrease the uncertainty interval by a factor of k requires increasing the number of independent observations by k^2. Thus, to gain the same improvement as the $3\times$ amplification would require collecting $9\times$ more data.

The most important aspect of Fig. 11 is that it illustrates that the accuracy with which a cancelling modulation can be estimated is jointly determined by the amplification factor, i.e. by the sensitivity of the method and by the amount of data collected. To sketch this more quantitatively, consider the points $a=(x_1, y_1)$ and $a'=(x_2, y_2)$, and assume equal intervals of uncertainty around each $\pm\sigma_\mu$. From Fig. 11, assuming a linear psychometric function in the range under consideration, it follows that the interval of uncertainty Δx around the 50% crossing is

$$\Delta x = \frac{\pm(\sigma_\mu)(x_2 - x_1)}{y_2 - y_1} \quad (1)$$

In a typical case, y_1, y_2 are 25 and 75%, respectively; $x_2 - x_1 = 2\text{JND}$, where JND represents the threshold for 25% and 75% correct motion direction responses; and $\sigma_\mu = \sqrt{y(1-y)/n}$ where n is the number of trials per point. Substituting in Eq. (1) yields $\Delta x = \pm 1.73\text{JND})/\sqrt{n}$. For $n=25$, this yields $\Delta x = \pm 0.346\text{JND}$. Thus, according to Eq. (1) (which is an approximation), a psychometric function with 25 trials per point reduces the uncertainty in the magnitude of a contaminating motion component by an additional factor of about 1/3 beyond what was achieved by motion amplification.

In theory, following the n^2 law for reducing the standard error of the mean by increasing the number n of trials, there is no limit on accuracy. In practice, it is easier to amplify, and a motion amplification factor of 3 or 4 is often attainable. Reducing the required n for a given level of accuracy by a factor of 9 or 16 by motion amplification is something like the difference between walking and driving. It makes a difference in the ventures that are undertaken.

[10] We have not considered explicitly here how second- and third-order components can be removed to create pure first-order stimuli. With luminance sinewaves, this is accomplished by adding stationary sine pedestal of $2\times$ the amplitude or by viewing very low contrast stimuli (for details, see Lu & Sperling, 1995a).

3.2. Comparison of sandwich amplification and minimum motion

Both of these methods offer the possibility of cancelling unwanted components in motion displays. The sandwich method (e.g. Fig. 1b and c) offers amplification by using large-modulation stimuli in the even (amplifier) frames, while the small amplitude stimuli are confined to the odd frames. The minimum motion technique (e.g. Fig. 1a–b,p–q) is restricted to displays that are near motion threshold, and it is inherently less sensitive. From the kinds of data obtained by the two methods, as schematically illustrated in Fig. 2, it is obvious that the estimate of the cancelling modulation m_c derives from the steepest part of the psychometric function, which yields the most accurate estimate. In the minimum motion method (Fig. 2b), estimates of the minimum where derivative is zero and where the second-derivative has its maximal value are inherently unreliable. Because of this, it is better but risky to assume symmetry of the data, and estimate the minimum of the minimum-motion curve from the midpoint between the steep side sections. All these considerations greatly favor the sandwich-amplification method.

On the other hand, the mathematics of sandwich-amplification yields simple results only when there are 90° frame-to-frame phase shifts between components being tested. And sandwich 'purification' applies precisely only when the resulting stimuli ultimately are used with 90° shifts. For example, a stimulus calibrated in a 90° sandwich might contain a 2nd harmonic which does not influence 90° motion (because its phase shift would be 180°). The 2nd harmonic would become relevant if the stimulus were moved in 45° steps. The minimum motion method can be used with any frame-to-frame phase shift or even continuous motion. Finally, minimum motion has face validity when the display actually used is the one being tested. So, a recommended procedure is to find an accurate m_c with a sandwich method and to check it with minimum motion. Again, the minimum motion check is possible only when the motion of interest is near its own motion-direction threshold.

4. Conclusion

Intensity distortion in early visual processing transforms stimuli so that, when they reach higher level visual processes, they contain undesired distortion products. Sandwich displays that exploit the amplification principle in motion perception to provide efficient calibration procedures for eliminating such undesired motion components. The resulting stimuli are certifiably pure in the sense that one can state that with probability P (e.g. 0.95) the stimuli contain less than some maximum amount (typically an order of magnitude less than threshold) of an undesired component. Methods are illustrated for producing 'pure' stimuli that individually stimulate the first-, second-, and third-order motion systems. Certifiably pure motion stimuli are a powerful tool, more generally, for revealing the properties of individual perceptual systems. Examples are given of contrast amplification in texture slant detection, and of the delicate assay of the effects of selective attention to perceptual features.

Acknowledgements

This research was supported by AFOSR, Life Science Directorate, Visual Information Processing Program. The authors thank P. Cavanagh for helpful comments.

References

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

Adelson, E. H. & Bergen, J. R. (1986). The extraction of spatio-temporal energy in human and machine vision, in motion: representation and analysis. IEEE Workshop Proceedings, IEEE Computer Society Press, Washington DC, pp. 151–155.

Allik, J., & Pulver, A. (1995). *Journal of Optical Society of America. A*, *12*, 1185–1197.

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

Anstis, S. M., & Cavanagh, P. (1983). A minimum motion technique for judging equiluminance. In J. D. Mollon, & E. T. Sharpe, *Colour vision* (pp. 155–166). New York: Academic Press.

Baylor, D. A., Hodgkin, A. L., & Lamb, T. D. (1974). Reconstruction of the electrical responses of turtle cones to flashes and steps of light. *Journal of Physiology (London)*, *242*, 759–791.

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

Boynton, R. M., Ikeda, M., & Stiles, W. S. (1964). Interactions among chromatic mechanisms as inferred from positive and negative incremental thresholds. *Vision Research*, *4*, 87–117.

Brainard, D. (1986). Cone contrast and opponent modulation color spaces. In: P. K. Kaiser & R.M. Boynton (Eds.), *Human color vision*, Optical Society of America, Washington DC, pp. 563–579.

Brainard, D. & Pelli, D. G. (1998). Psych Toolbox for Macintosh Computers, 1998.

Cavanagh, P., & Anstis, S. (1991). The contribution of color to motion in normal and color-deficient observers. *Vision Research*, *31*, 2109–2148.

Chichilnisky, E.-J., Heeger, D., & Wandell, B. A. (1993). Functional segregation of color and motion perception examined in motion nulling. *Vision Research*, *33*, 2113–2125.

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

Chubb, C., & Sperling, G. (1989b). Second-order motion perception: space–time separable mechanisms. In *Proceedings*: *Workshop on Visual Motion* (*March* 20–22, 1989, *Irvine*, *CA*) (pp. 126–138). Washington DC: IEEE Computer Society Press.

Chubb, C., Lu, Z.-L., & Sperling, G. (1999). Measuring the nonlinearity used to sense high temporal frequency second order motion. *Investigative Ophthalmology and visual Science*, *ARVO Supplement*, *40*, S424.

Derrington, A. M., Krauskopf, J., & Lennie, P. (1984). Chromatic mechanisms in lateral geniculate nucleus of macaque. *Journal of Physiology (London)*, *357*, 241–265.

Knutsson, H. & Granlund, G. H. (1983). Texture analysis using two-dimensional quadrature filters. *1983 IEEE Computer Society Workshop on Computer Architecture for Pattern Analysis and Image Database Management*, CAPAIDM, Silver Spring, Maryland: IEE Computer Society, pp. 206–213.

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

Lu, Z.-L., & Sperling, G. (1995b). Attention-generated apparent motion. *Nature*, *379*, 237–239.

Lu, Z.-L., & Sperling, G. (1996a). Three systems for visual motion perception. *Current Directions in Psychological Science*, *5*, 44–53.

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

Lu, Z.-L., & Sperling, G. (1999a). The amplification principle in motion perception. *Investigative Ophthalmology and visual Science*, *ARVO Supplement*, *40*(4), S199.

Lu, Z.-L., & Sperling, G. (1999b). Unequal representation of black and white in human vision. *Investigative Ophthalmology and visual Science*, *40*(4), S200.

Lu, Z.- L. & Sperling, G. (2001). The three systems theory of human visual motion perception: review and update. *Journal of the Optical Society of America*, A, submitted.

Lu, Z.-L., Lesmes, L., & Sperling, G. (1999). Mechanisms of isoluminant chromatic motion perception. *Proceedings of National Academy of Sciences USA*, *96*, 8289–8294.

He, S., & MacLeod, D. I. A. (1998). Contrast-modulation flicker: dynamics and spatial resolution of the light adaptation process. *Vision Research*, *38*, 985–1000.

Moreland, J. D. (1980). A modified anomaloscope using optokinetic nystagmus to define colour matches objectively. In G. Verriest, *Colour vision deficiencies* (pp. 189–191). Bristol UK: Hilger.

Morgan, M. J., & Chubb, C. (1999). Contrast facilitation in motion detection: evidence for a Reichardt detector in human vision. *Vision Research*, *39*, 4217–4231.

Nakayama, K., & Silverman, G. H. (1985). Detection and discrimination of sinusoidal grating displacements. *Journal of the Optical Society of America*, *A*, *2*, 267–274.

Pelli, D. G., & Zhang, L. (1991). Accurate control of contrast on microcomputer displays. *Vision Research*, *31*, 1337–1350.

Prinzmetal, W., Amiri, H., Allen, K., & Edwards, T. (1998). Phenomenology of attention: I. Color, location, orientation, and spatial frequency. *Journal of Experimental Psychology: Human Perception and Performance*, *24*, 261–282.

Reichardt, W. (1957). Autokorrelationsauswertung als funktionsprinzip des zentralnervensystems. *Zeitschrzft Natuiforschung*, *12b*, 447–457.

Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In W. A. Rosenblith, *Sensory communication* (p. 1961). New York: Wiley.

Scott-Samuel, N. E., & Georgeson, M. A. (1999). Does early non-linearity account for second-order motion? *Vision Research*, *39*, 2853–2865.

Smith, A. T., & Ledgeway, T. (1997). Separate detection of moving luminance and contrast modulations: fact or artifact? *Vision Research*, *37*, 45–62.

Solomon, J. A., & Sperling, G. (1994). Full-wave and half-wave rectification in 2nd-order motion perception. *Vision Research*, *34*, 2239–2257.

Sperling, G., & Lu, Z.-L. (1998a). A systems analysis of motion perception. In T. Watanabe, *High-level motion processing* (pp. 153–183). Cambridge, MA: MIT Press.

Sperling, G., & Lu, Z.-L. (1998b). Update on the status of the three-motion-systems theory. *Investigative Ophthalmology and visual Science*, *ARVO Supplement*, *39*, 461s.

Stromeyer, C. F., Kronauer, R. E., Ryu, A., et al. (1995). Contributions of human long-wave and middle-wave cones to motion detection. *Journal of Physiology (London)*, *485*, 221–243.

Taub, E., Victor, J. D., & Conte, M. M. (1997). Nonlinear preprocessing in short-range motion. *Vision Research*, *37*, 1459–1477.

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

van Santen, J. P. H., & Sperling, G. (1985). Elaborated Riechardt detectors. *Journal of the Optical Society of America A: optics and image science*, *2*, 300–321.

Watson, A. B., & Ahumada, A. J. Jr (1983). A look at motion in the frequency domain. In J. K. Tsotsos, *Motion*: *perception and representation* (pp. 1–10). New York: Association for Computing Machinery.

Werkhoven, P., Sperling, G., & Chubb, C. (1993). Motion perception between dissimilar gratings: a single channel theory. *Vision Research*, *33*, 463–465.