Perception & Psychophysics
1998, 60 (4), 558-574

Opponent motion interactions in the perception of transparent motion

DELWIN T. LINDSEY
Ohio State University, Columbus, Ohio
and Ohio State University, Mansfield, Ohio

and

JAMES T. TODD
Ohio State University, Columbus, Ohio

Interactions in the perception of motion transparency were investigated using a signal-detection paradigm. The stimuli were the linear sum of two independent, moving, random-check "signal" textures and a third texture consisting of dynamic random "noise." Performance was measured as the ratio of squared signal and noise contrasts was varied (S^2/N^2). Motion detectability was poorest when the two signal textures moved in opposite directions (180°), intermediate when they moved in the same direction (0°), and best when the textures moved in directions separated by 90° in the stimulus plane. This pattern of results held across substantial variations in velocity, field size, duration, and texture-element size. Motion identification was also impaired, relative to 0°, in the 180° but not in the 90° condition. These results are consistent with the idea that performance in the opponent-motion condition is limited by inhibitory (or suppressive) interactions. These interactions, however, appear to be direction specific: little, if any, inhibition was observed for perpendicular motion.

Motion transparency refers to the perception of two or more directions of motion that are present simultaneously at the same location in visual space. The earliest systematic studies of this phenomenon were primarily concerned with the kinetic depth effect (Wallach & O'Connell, 1953); in particular, they focused on the conditions for specifying depth in moving 2-D stimuli (e.g., Balch & Shaw, 1978; Gibson, Gibson, Smith, & Flock, 1959; Mace & Shaw, 1974). For example, Gibson et al. presented subjects with two sets of random dots that were uniformly translated through one another. The subjects not only perceived motion simultaneously in the directions of translation of the two stimulus components, but also perceived a separation of the two components in depth, so that the two sets of dots appeared as two surfaces, with one moving "in front of" the other. Subsequent studies have demonstrated that up to three surfaces can be distinguished when presented simultaneously in motion at different velocities (e.g., Andersen, 1989; see also Andersen & Wuestefeld, 1993). While these studies were concerned primarily with the relationship between motion and the specification of depth in 2-D stimuli, the phenomenon of motion transparency has become useful in the study of the general problem of computing 2-D object motion from patterns of optical flow to which low-level motion sensors respond. In the experiments described below, we used motion transparency as a means of exploring the role played by early inhibitory or suppressive interactions in these computations.

The idea that inhibitory interactions occur early in visual motion processing was originally associated with the view that low-level motion processing was inherently opponent in nature. According to this view, an early representation of optical flow is derived from subtractive interactions between motion sensors that differ from one another by 180° in the directions of motion to which they respond optimally. Reichardt (1961) incorporated the concept of motion opponency into his model of motion processing in insects. Motion opponency also has figured prominently in a number of subsequent models of motion sensors in humans (Adelson & Bergen, 1985; Barlow & Levick, 1965; van Santen & Sperling, 1984, 1985; Watson & Ahumada, 1985).

Motion opponency was an appealing concept because it accounted qualitatively for two kinds of phenomena in motion perception. The first was adaptation phenomena such as the motion aftereffect. If a subject was adapted to motion in one direction, then a subsequently viewed, stationary stimulus would appear to move in a direction opposite to that of the adapting stimulus (see Wohlgemuth, 1911). This suggested (1) that the perception of motion involved a balance between mechanisms tuned to opposite directions of motion, (2) that adaptation to motion in one

This research was supported by Grants F49620-93-0116, from the Air Force Office of Scientific Research, and SRB-9514522, from the National Science Foundation. Angela M. Brown provided valuable criticism on an earlier version of this manuscript. Summaries of Experiments 1 and 2 were previously reported at the 1994 and 1995 annual meetings of the Association for Research in Vision and Ophthalmology. Address correspondence concerning this article to D. T. Lindsey, Ovalwood Hall, Ohio State University, Mansfield, Ohio 44906 (e-mail: dlindsey@magnus.acs.ohio-state.edu).

direction would create an imbalance, and (3) that the imbalance was revealed when the stationary stimulus was viewed following adaptation. The second was the finding that counterphase-modulated cosine gratings, although they simultaneously generated stimuli for motion in two opposing directions, almost always elicited the perception of flicker rather than motion in human subjects. The perception of flicker was attributed to the mutual cancellation of motions of equivalent strength in the two opposing directions. Stromeyer, Kronauer, Madsen, and Klein (1984) have also shown that cosine gratings moving in opposite directions effectively cancel each other's detectability when superimposed on suprathreshold, counterphase-modulated masks.

Although the concept of motion opponency, as it was originally formulated, is able to account for some aspects of motion perception, a weakness has always been its inability to deal with phenomena related to motion transparency. For example, motion transparency is often seen in some bidirectional stimuli, even when the components are moving in opposite directions (see, e.g., Clarke, 1977; Gibson et al.,1959; van Doorn & Koenderink, 1982a). Furthermore, the motion aftereffect elicited by transparent bidirectional stimuli consisting of textures moving at right angles to one another is not bidirectional, as suggested by the concept of motion opponency, but is in a direction opposite to the average motions of the two transparent textures (Mather, 1980; van Doorn, Koenderink, & van de Grind, 1984, 1985; see also Verstraten, Fredericksen, & van de Grind, 1994).

More recently, inhibitory interactions have been incorporated into a number of multistage models of motion processing that focus explicitly on the computation of motion from the responses of low-level motion sensors. These models generally consist of the sensor stage, one or more pooling stages, and a decision stage. The sensor stage consists of 2-D arrays of motion sensors that are assumed to be spatially orientation selective and to respond best to motion at right angles to their preferred orientation. Pooling stages are needed because a moving object generally produces responses within a number of such sensor arrays which are tuned to respond optimally to a different range of spatial frequencies, orientations, speeds, and directions of motion. The purpose of the pooling and decision stages is to compute a single unified object velocity from the integrated responses of these different sensor populations, and to do so in a way that permits the spatial resolution of multiple objects in the visual field.

It is not known exactly how this might be done in a way that is compatible with motion transparency. Many early computational models of motion proposed a pooling stage that performed a smoothing or regularization operation that was based on certain assumptions about patterns of optical flow or on least squares methods involving the responses of motion sensors from several neighboring regions of visual space (e.g., Grzywacz & Yuille, 1990; Heeger, 1987; Hildreth, 1984; Horn & Schunk, 1981; Poggio, Yang, & Torre, 1988). These methods worked well for recovering estimates of motion in a single direction, but did not work at all for recovering transparent motion since, by design, they assumed that all responses within a region of visual space arose from sensor responses due to a single moving object, noise, or a combination of the two. More recent computational approaches to the recovery of object motion and motion transparency have proposed pooling processes that permit the parallel evaluation of motion in multiple velocities (e.g., Jasinschi, Rosenfeld, & Sumi, 1992; Kim & Wilson, 1992; Nowlan & Sejnowski, 1995; Qian, Andersen, & Adelson, 1994b). The focus of these new approaches has been on the methods by which local sensor measurements arising from two or more transparent layers can be appropriately partitioned for analysis.

A particularly intriguing approach to this problem is the one suggested by Qian, Andersen, & Adelson (1994a, 1994b). Their proposal seeks to account for why some stimuli consisting of two moving components elicit the perception of motion transparency while others do not, even though both kinds of stimuli should, in principle, generate strong responses in motion sensors tuned to respond to the two moving components. A key aspect of their proposal is the idea that, early in motion processing, mutual inhibitory or suppressive interactions occur among the responses of sensors tuned to different directions of motion, while mutually facilitating interactions occur among sensors tuned to similar directions of motion. These early interactions are hypothesized to occur within a very narrow window of time and within regions of space corresponding to the sizes of the receptive fields of the motion sensors themselves. According to Qian et al., counterphase-modulated gratings do not elicit motion transparency, because the motion signals generated by the two components are locally balanced—that is, of equal magnitude everywhere in the stimulus field—and effectively cancel one another. By contrast, when two interdigited random-dot patterns are translated in opposite directions, the motion signals elicited by these two patterns vary stochastically with location across the stimulus field and the cancellation of motion is incomplete.

The present series of studies had three goals. The first was to measure the effects of putative inhibitory interactions in the processing of transparent motion stimuli and to see whether there was a directional bias in these interactions. The second goal was to explore the generality of our findings across a wide variety of potentially important stimulus parameters. And the third goal was to relate our findings to current theories regarding the computation of 2-D motion. Rather than focusing on the specific computation of the 2-D speeds and directions of objects, we approached our study of motion transparency from the perspective of signal detection theory. That is, we measured human observers' abilities to detect and identify motion that was embedded in spatially coincident dynamic random noise.

The experimental paradigm that we employed was similar to the one devised by van Doorn and Koenderink

(e.g., 1982a, 1982b). In our version of this paradigm, the stimuli were determined by the spatiotemporal characteristics of three independently computed, spatially superimposed, additively combined dynamic random-texture patterns. The luminance of a pixel in the final stimulus texture displayed on the monitor corresponded to the sum of the corresponding pixel luminances across the three component textures. The elements of two of these textures were translated an integer number of elements per unit time, either in the same direction or in different directions. These were the two "signal" layers of the stimulus. The third pattern, referred to as the "noise" layer, was generated anew on each frame of the animation sequence and therefore did not contain any net motion signal. We were interested in how well subjects would perform in detecting or identifying motion in mixtures of signal and noise as their ratio—the signal-to-noise ratio, or SNR—was decreased.

We chose a signal-detection paradigm because putative inhibitory or suppressive interactions in early motion processing should lead to some degree of cancellation of the neural responses that subjects may monitor in a signal-detection task. Accordingly, signal cancellation will reduce subjects' sensitivities, thus making it more difficult for them to distinguish between a stimulus consisting of signal plus noise and a stimulus consisting of pure noise at a given value of SNR. The signal-detection paradigm does not implicitly distinguish between signal cancellation and other potential forms of information loss. Therefore, the design of the experiments described below permitted comparisons of performance when the two signal layers were moving in the same direction with performances when the signal layers were moving in different directions. In interpreting our results, we assume that information losses due to factors other than signal cancellation remain constant as the relative directions of motion of the signal layers are manipulated.

GENERAL METHOD

Experimental Design

The study consisted of three experiments. In Experiment 1, we used a two-interval forced-choice technique and obtained psychometric functions for the discrimination between a stimulus consisting of some mixture of signal and noise and a stimulus consisting of noise alone. The psychometric functions were obtained for a suite of five motion conditions. In three of these conditions, the two signal layers both moved coherently in one of the following three directions: left, right, or down. We will refer to these conditions variously as the *nontransparent* or *unidirectional* motion conditions. In the other two conditions, which we will refer to throughout this paper as the *transparent* or *bidirectional* motion conditions, the signal layers moved at the same speed but in different directions. In one case, the *opponent* motion condition, the stimuli consisted of superimposed leftward- and rightward-moving patterns. In the fifth and final condition, referred to as the *perpendicular* motion condition, leftward- and downward-moving patterns were combined. This suite of five conditions allowed us to make two comparisons of detectability that were fundamental to our study. First, we could compare the effects of one versus two sources of motion—unidirectional versus bidirectional—on the detectability of motion in our dynamic random-check patterns. Second, we could determine the degree to which the relative principal directions of the motion signals—opponent versus nonopponent—influence the detectability of motion.

In Experiment 2, we examined the generality of the findings of Experiment 1 across a variety of different physical stimulus parameters. While we obtained the data in Experiment 1 with a particular speed, field size, check size, and stimulus duration, in Experiment 2, we measured motion detection thresholds as functions of each of these variables. The comparison was made across the same suite of five motion conditions employed in Experiment 1.

In Experiment 3, we examined the generality of the findings of Experiment 1 across tasks. While Experiments 1 and 2 were designed to measure motion detectability, in Experiment 3, we measured threshold SNRs for identification of directions of motion in dynamic random-check stimuli.

Subjects

The first author served as a subject in all of the experiments described below. One or more additional subjects were recruited for each of the experiments. All subjects were either emmetropic or had normal visual acuity with appropriate corrective lenses.

Apparatus

The animation sequences were generated on an SGI VXGT color graphics workstation (Silicon Graphics, Inc.) and displayed on a 20-in. high-resolution RGB monitor (1,280 horizontal × 1,024 vertical pixels; 24 bits/pixel; 60 Hz noninterlaced). At the viewing distance of 57.3 cm, each pixel subtended 1.59′ of arc visual angle. A mean stimulus luminance of 97 cd/m^2 was maintained throughout the study. The interval between the onsets of frames in the animation sequences was always 1/60 sec. The SGI workstation employs a double-buffering scheme, and the transition from one presentation to the next therefore occurred during the display's vertical-blanking interval.

The computer monitor was corrected for nonlinearities in its light output by measuring relative radiance with a PIN10 (United Detector Technologies) photodetector as a function of pixel values for the red, green, and blue guns of the monitor. The measurements were then digitized and averaged. This calibration matrix was then inverted to produce the desired correction vector for the video display controller. An additional series of measurements, designed to assess the accuracy of the relative radiance calibration, was then undertaken at regular intervals during the course of the study. The radiances of individual elements of very coarse random-check patterns were measured as a function of the expected SNR. From these values, an "actual" SNR was calculated. Typically, the expected and actual SNRs differed from one another by 2% or less and never by more than 10%.

Stimuli

The stimuli consisted of dynamic superpositions of three scaled pseudorandom textures, $\mathbf{P}_k$. The elements of $\mathbf{P}_k$ could assume one of two values: −1 or 1. $\mathbf{P}_1$ and $\mathbf{P}_2$, were the "signal" textures, and $\mathbf{P}_3$ was the "noise" texture. Let $\bar{L}$ represent the mean luminance of the three-texture composite stimulus. Then the luminance of the ijth visible check in the tth frame of the animation sequence was specified in the following way:

$$L_{\text{Stim}}(i,j,t) = \bar{L} + q\{\mathbf{P}_1(i - v_{1i}t, j - v_{1j}t) + \mathbf{P}_2(i - v_{2i}t, j - v_{2j}t)\} + r\,\mathbf{P}_3(i,j,t), \quad (1)$$

where $\{v_{1i}, v_{1j}\}, \{v_{2i}, v_{2j}\}$ correspond to the displacements of the elements in the two signal textures, $\mathbf{P}_1$ and $\mathbf{P}_2$, from one frame of the animation sequence to the next; q and r are scalars that determine the relative proportions of $\mathbf{P}_1$ and $\mathbf{P}_2$ and of $\mathbf{P}_3$, and therefore

the proportions of signal and noise, in the dynamic stimulus; $L_{Stim}(i,j,t)$ is mapped to the central square region of an otherwise dark computer monitor. In all but one series of measurements described below, each check consisted of an individual pixel subtending 0.027° × 0.027° of visual angle.

Figure 1 depicts two successive frames of animation sequences in which the two signal layers are either both moving in the same direction (left only) or in different directions (left + down). Panel a depicts a magnified 16 × 16 pixel portion of a stimulus at frame n of either animation sequence. It consists of an appropriately scaled sum of the binary random-check textures depicted in panels b–d. Panels b and c depict the structures of the two signal textures, S_1 and S_2. These patterns consist of portions of binary patterns $\mathbf{P}_1$ and $\mathbf{P}_2$, scaled by q. Panel d depicts the noise layer N, consisting of the pattern $\mathbf{P}_3$, scaled by r. The middle column of Figure 1 (panels e–h) depicts the same 16 × 16 pixel portion of a stimulus and its three components—S_1, S_2, and N—at frame $n + 1$. Here the composite pattern seen by the subject (panel e) consists of two signal patterns, S_1 and S_2 (panels f and g), which have each moved 1 pixel to the left of their respective positions in the previous frame, and a noise layer N (panel h), which is uncorrelated with the noise layer in the previous frame. The rightmost column (panels i–l) depicts the compos-

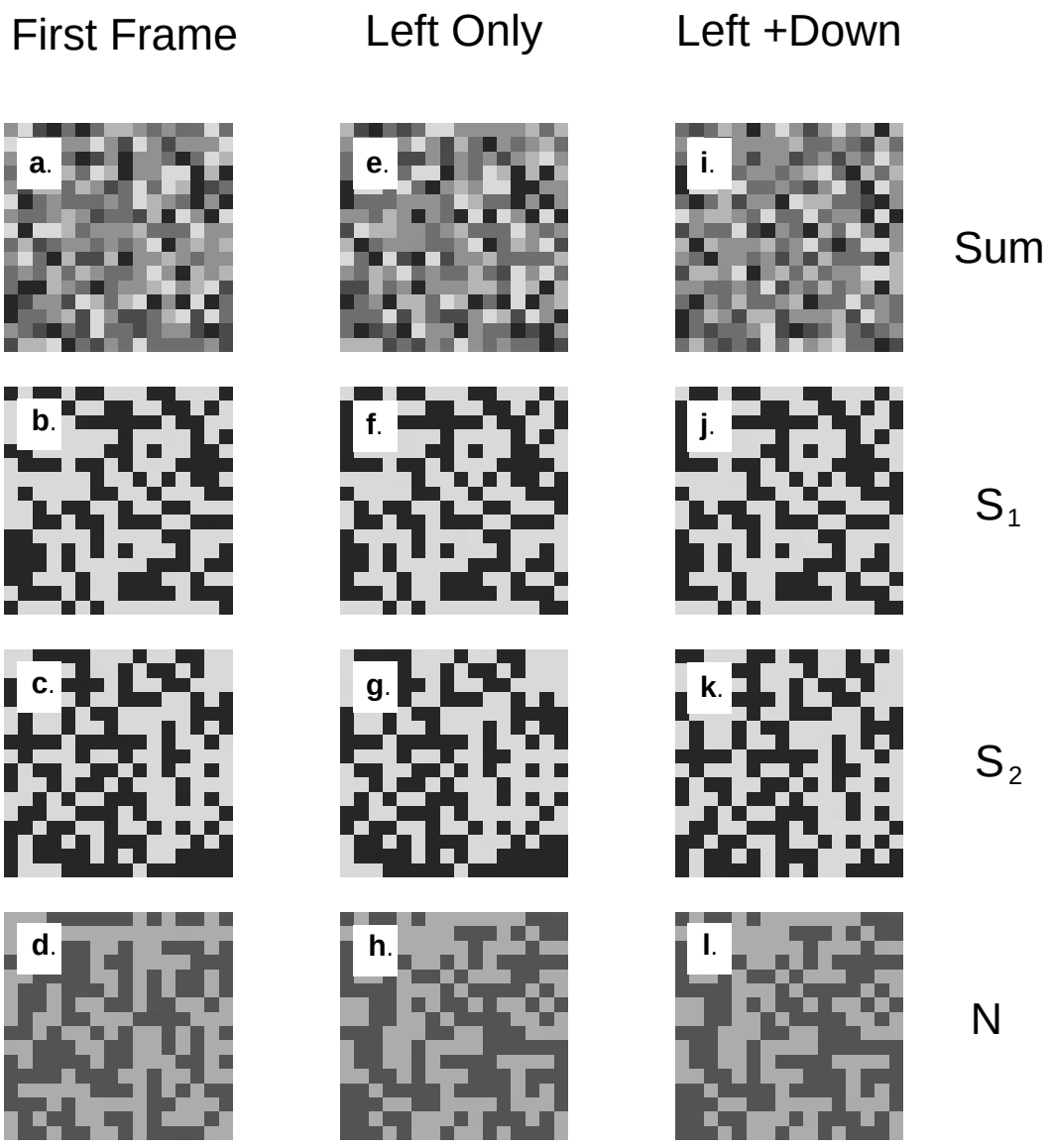

Figure 1. Magnified 16 × 16 pixel portions of two two-frame sequences of dynamic random-check textures, illustrating method of producing unidirectional and bidirectional stimuli. Left column: One frame in an animation sequence. The random-check texture (a) seen by a subject consists of the sum of two signal patterns, S_1 (b) and S_2 (c), and a noise pattern, N (d). Middle column: Same region of the display during the next frame of an animation sequence. Both signal textures have moved 1 pixel to the left (compare b and f; c and g). The structure of the noise texture is independent of that in the preceding frame (compare d and h). Right column: Same as middle column, except that S_1 has moved leftward 1 pixel (compare b and j); S_2 has moved downward 1 pixel (compare c and k).

ite and its components in frame $n + 1$ of a sequence in which S_1 is moving to the left and S_2 is moving downward 1 pixel per frame.

The Specification of SNR

Following van Doorn and Koenderink (1982a), the signal and noise components of the stimulus were specified by the relative contributions of the three pseudorandom binary arrays, $\mathbf{P}_k$, to VAR(L_{Stim}), the squared RMS contrast of the stimulus. VAR(L_{Stim}) was always 0.281, which corresponded to an RMS contrast of 0.53. The relative values of q and r were manipulated to maintain this value of RMS contrast while varying the proportions of signal and noise contributed to the stimulus by $\mathbf{P}_k$. Since the value assigned to a given element of any array, $\mathbf{P}_k$, was an independent random variable, the luminance variance of the stimulus was equal to the sum of the weighted variances contributed by $\mathbf{P}_1$ and $\mathbf{P}_2$ and $\mathbf{P}_3$:

$$\text{VAR}(L_{\text{Stim}}) = q^2\{\text{VAR}(\mathbf{P}_1) + \text{VAR}(\mathbf{P}_2)\} + r^2\text{VAR}(\mathbf{P}_3); \quad (2)$$

$$\text{VAR}(\mathbf{P}_k) = \frac{\sum \mathbf{P}^2_{i,j,t}}{n}, \quad (3)$$

where $\mathbf{P}_{i,j,t}$ is the value of the ijth pixel in the tth animation frame and n is the total number of pixels over which the variance is computed.

In all of the experiments described below, SNR is reported as the ratio between the luminance variance associated with motion in any given direction and the remaining luminance variance.[1] In the case of the unidirectional motion conditions, both $\mathbf{P}_1$ and $\mathbf{P}_2$ contribute to the computation of stimulus signal. The variances of $\mathbf{P}_1$ and $\mathbf{P}_2$ add, and SNR is computed in the following way:

$$\text{SNR} = \frac{\text{Signal}_{\text{Stim}}}{\text{Noise}_{\text{Stim}}} = \frac{q^2\{\text{VAR}(\mathbf{P}_1) + \text{VAR}(\mathbf{P}_2)\}}{r^2\text{VAR}(\mathbf{P}_3)} = \frac{2q^2}{r^2}. \quad (4)$$

The specification of SNR is not straightforward in the case of the transparent motion conditions, since the signal layers move in different directions and will, therefore, be processed by different populations of motion sensors. We chose to specify bidirectional SNR as the ratio of the luminance variance *in either signal layer* and the luminance variance in the noise layer[2]:

$$\text{SNR} = \frac{q^2\text{VAR}(\mathbf{P}_{1\text{or}2})}{r^2\text{VAR}(\mathbf{P}_3)} = \frac{q^2}{r^2}. \quad (5)$$

In comparing the results for bidirectional and unidirectional motion conditions in the experiments described below, it is important to recognize that the bidirectional motion conditions provide two signal sources of equal signal level at a given value of SNR while the unidirectional condition provides only one signal source. The effects of multiple signal sources on a subject's performance will depend on the nature of the visual task. We expect little or no effect of multiple signal sources in our motion-identification task, because decisions must be based on identification of a single, specified direction of motion. However, we expect the presence of multiple signal sources to facilitate detection, since subjects are required only to detect the presence of motion and can therefore base their decisions in our detection task on visual responses elicited by either or both of these signal sources. The expected magnitude of this facilitation will depend on many factors, including (1) the assumed properties of the visual mechanisms that underlie motion detection, (2) sources of variance in detector responses, (3) uncertainty effects introduced by the design of the forced-choice task, and (4) the decision rule employed by the subjects (see Graham, 1989, for a comprehensive discussion). In the absence of suppressive or inhibitory effects, a reasonable upper-bound estimate of the facilitating effects of multiple signal sources based on signal detection theory is that detectability, d', for the bidirectional conditions should be equal to the sum of the d's for the individual components.[3] Thus, for equally detectable components in a bidirectional stimulus, d' should be no greater than twice either of the component d's at a given SNR level. Differences between the unidirectional and bidirectional conditions in threshold SNR will depend upon the choice of a criterion d' and upon the sensitivities of the subjects to the unidirectional motion components. Furthermore, if suppressive interactions among sensors tuned to different directions of motion exist, detectability in the bidirectional cases should be less than predicted by probability summation.

While interpretation of the results for the uni- and bidirectional conditions may depend to some extent on a quantitative model of signal detectability, comparison of the results for the opponent and perpendicular bidirectional conditions does not. They differ only in the direction of motion of one of their components (right vs. down) and should yield similar detectabilities unless there exist significant anisotropies in either motion detectability, per se, or in the effects of suppressive interactions across sensors tuned to different directions of motion. The measurment of uni- and bidirectional detectabilities in the same experimental session helps decide between these two alternatives.

General Procedure

The subject, who sat in front of the computer monitor in a room illuminated only by the light from the monitor, adapted to a uniform field of the same size and mean luminance as that of the stimulus for 1 or 2 min prior to the start of data collection. The viewing distance was maintained and head movements were minimized by using a conventional chin/head-restraint system. The subject fixated a small stationary red spot in the center of the stimulus display region of the RGB monitor throughout a trial. The SGI's three-button mouse was used by the subject to initiate each trial and then to record his/her response.

We employed a two-interval forced-choice procedure (2-IFC) in Experiments 1 and 2. Each trial consisted of two stimulus intervals. In one pseudorandomly chosen interval, the stimulus consisted of a mixture of signal and noise; in the other interval, the stimulus consisted entirely of noise. The stimuli presented in each of the two intervals had identical values of q and r and differed from one another only in the temporal structures of $\mathbf{P}_1$ and $\mathbf{P}_2$, which were coherent in the signal-plus-noise interval and pseudorandom in the noise-alone interval. At the end of each trial, the subject pressed a button indicating the interval containing the signal-plus-noise stimulus. Additional features of the forced-choice procedures used in the experiments are described in the appropriate sections below.

Forced-choice techniques were also used in Experiment 3 to estimate identification thresholds. These techniques are described in detail in the appropriate section below.

EXPERIMENT 1

In Experiment 1, psychometric functions were measured for discriminating mixtures of signal and noise from pure noise. Performance was assessed separately for each of five motion signal conditions: left alone, right alone, down alone, left plus right, and down plus left. The first three conditions are the *component* motion conditions; the last two constitute the *bidirectional* motion conditions, opponent and nonopponent, respectively. Recall that the effects of multiple motion sources on detectability can be evaluated by comparing unidirectional and bidirectional motion-detection performance, whereas potentially direction-specific effects of these multiple sources can be evaluated by comparing performance on the two bidirectional motion conditions.

Method

The dynamic random-check stimuli were presented within a square aperture subtending 6.8º (256 pixels) on a side. The average speed of each signal layer was always 12.7º/sec (8 pixels per frame). The signal-plus-noise and noise-alone stimulus intervals were obtained by multiplying each animation sequence by a trapezoidal temporal contrast window. The rising and falling phases of each window were both 30 frames (0.5 sec) in duration. The plateau of the waveform, corresponding to an RMS contrast of 0.53, had a duration of 100 frames (1.67 sec). A period of 0.5 sec separated the end of the first stimulus interval and the beginning of the second interval. Thus, the total time required to present both stimulus intervals in a given trial was 3.84 sec.

Subject performance was measured using the method of constant stimuli coupled with the 2-IFC procedure described in the General Methods section above. Percent correct responses were measured at 4 or 5 logarithmically spaced SNR levels per motion condition. The SNRs were chosen for each subject on the basis of extensive practice sessions. Data for a particular subject were usually gathered over the course of several experimental sessions. Stimuli were presented in block-random fashion, where the size of each block of trials was 5 motion conditions × 4 SNR levels = 20 conditions per block. A typical experimental session consisted of 20 such blocks of trials, and usually lasted approximately 40 min. The results reported below are based on 100 trials per condition.

Results

The results of Experiment 1 are shown in Figure 2. The ordinate plots values of detectability, d', obtained by converting each fraction-correct response obtained with the forced-choice procedure into the corresponding d', using the standard relationship that assumes that subjects' hit and false-alarm rates are equivalent:

$$d' = \sqrt{0.5} \cdot z(f), \tag{6}$$

where $z(f)$ is the normal deviate corresponding to the percent correct, f (Green & Swets, 1966; see also MacMillan & Creelman, 1991). The abscissa values indicate relative proportions of signal in the mixtures of signal and noise being detected.[4] Filled and unfilled symbols indicate d's for bidirectional and unidirectional conditions, respectively. The solid lines are least squares fits to the opponent and perpendicular motion data, while the dotted line is fitted to the averaged unidirectional motion data. The dashed line in each panel indicates the upper bound estimate for probability summation, as discussed in the General Method section above. This line has a slope twice that based on the averaged unidirectional data.

In signal-detection analysis, sensitivity is indicated by the rate of change of d' as signal level is increased. All 3 subjects were most sensitive to the down-left, or nonopponent, transparent-motion condition and, on average, least sensitive to the left-right, or opponent, motion condition. On the average, sensitivities for the perpendicular- and opponent-motion conditions differed by a factor of 2.01. Differences in performance for the two transparent-motion stimuli can also be expressed as threshold SNRs, the SNR corresponding to a d' of 0.96, or 75% correct in a 2-IFC experiment. In that case, thresholds for the perpendicular-motion conditions were, on average, 1.65 times lower than those for opponent motion. Performance in the case of unidirectional motion tended to fall intermediately between the two extremes.

Several features of the data stand out when performance is compared either in terms of the slopes of the signal-detection curves or in terms of threshold-level performance. One feature is that the differences in performance

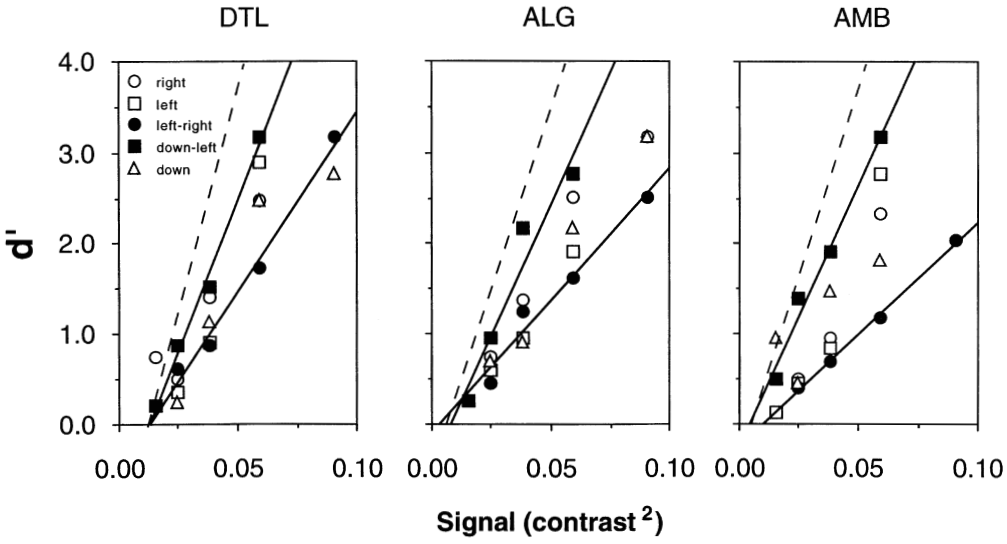

Figure 2. Detectability, d', of motion in unidirectional and bidirectional random-check textures as a function of the motion signal level. Data for 3 subjects are shown. Each panel shows data for the standard suite of five signal layer conditions: right only, left only, left right, down left, and down only. See figure legend for identification of plotting symbols. The solid lines in each panel are linear fits to the left-right, opponent-motion condition (lower solid line) and the down-left, perpendicular-motion condition (upper solid line). The dashed line in each panel corresponds to a line with a slope twice that derived from a linear fit to the average data obtained for the three unidirectional motion conditions in the suite. This represents an upper-bound estimate for the effects of probability summation on motion detection in the bidirectional conditions.

among the different conditions examined were modest, amounting to no more than a factor of about 2. A second feature of the data is that the values of d' for opponent motion in each panel fall systematically below the corresponding values of d' for unidirectional motion. These results clearly indicate the effects of signal suppression in the opponent motion condition. We believe that were it not for probability summation these suppressive effects would be even greater than indicated in Figure 2.

Another feature of the data is that d' rose most rapidly with signal level in the perpendicular-motion condition across all 3 subjects tested. These data are consistent with probability summation of visual responses evoked by the two signal sources, leftward and downward in this condition. Note, however, that the degree of summation is less than the upper-bound indicated in the panel for each subject. This may indicate that, in addition to probability summation, some degree of signal suppression or inhibition occurred in the perpendicular conditions. Another obvious possibility is that the probability summation actually exhibited by the human visual system may not follow the signal summation rule described in the General Methods section. In any event, the results of Experiment 1 demonstrate that, whatever the decision model, the factors that govern signal detectability in the two bidirectional-motion conditions are direction specific: although the two bidirectional-motion conditions differ only in the direction of motion of one of the stimulus components (downward vs. rightward), performance is lower for opponent than for perpendicular motion. Furthermore, subjects' performance averaged a factor of about 1.5 less on the opponent-motion condition than on unidirectional-motion conditions, in spite of the potential benefits of probability summation.

EXPERIMENT 2

In Experiment 2, we measured threshold SNRs for detecting motion produced by the suite of five signal conditions used in Experiment 1, as functions of the stimulus speed, field size, check size, and duration. Experiment 2 was designed to explore the generality of the results obtained in the preceding experiment. Were the modest direction-specific differences in performance observed in Experiment 1 sustained across a wide variation in stimulus parameters or were the conditions under which Experiment 1 was performed special in some fundamental way? The effects of many of these variables on motion detection had been examined extensively using dynamic random-check patterns consisting of a single motion layer (e.g., Koenderink, van Doorn, & van de Grind, 1985; van de Grind, Koenderink, & van Doorn, 1986; van de Grind, van Doorn, & Koenderink, 1983; van Doorn & Koenderink, 1982a, 1982b, 1984). However, the corresponding dependencies of motion detection in stimuli consisting of multiple motion layers had not been systematically investigated.

Method

Stimuli

Speed. Detection-threshold SNRs were measured for four values of average speeds spanning 1.59º to 12.8º/sec in 1-octave steps. Each desired average speed was obtained by displacing the random-check patterns 1, 2, 4, or 8 pixels per frame of the animation sequence.

Duration. Detection thresholds were measured as functions of stimulus duration at each of two speeds: 1.6º and 12.7º/sec. Durations varied from 83 to 1,667 msec. The speed, field size, and check size of the stimuli matched those employed in Experiment 1. Accurate timing of the stimulus interval dictated a rectangular rather than the trapezoidal temporal contrast envelope used in Experiment 1. To mitigate the potential effects of contrast masking at the onset of the animation sequences, the following presentation sequence was employed in each stimulus interval. At the beginning of a trial, the uniform gray pretrial stimulus field was replaced by a stationary texture pattern corresponding to the first frame of the stimulus-animation sequence. Following a period of 0.25 sec, all but the last of the remaining frames in the animation sequence for that interval of the trial were presented at 60 frames/sec. The last frame of the animation sequence remained visible in the stimulus field for 0.25 sec, and was then replaced by a uniform gray field. Following a 0.5-sec interstimulus interval, the presentation sequence was repeated for the second interval of the trial.

Field Size. The effects of field size on motion-detection threshold were investigated in two series of experiments. In one series of measurements, detection-threshold SNRs were measured for field sizes of 2.88 to 46.10 deg^2 in 2-octave steps. Due to hardware limitations in our computer graphics system, thresholds for field sizes larger than 46.10 deg^2 could not be measured with single-pixel-per-check stimuli without reducing the animation rate. Therefore, measurements of detection threshold were repeated on 1 subject (D.T.L.), using a hardware zoom feature that increased check size from 1 to 8 pixels (0.21º) on a side and permitted animation of larger patterns at a 60-Hz frame rate. In this second series of measurements, field size ranged from 2.88 to 184.4 deg^2 in 2-octave steps. The velocity of the stimuli in both series of experiments was 12.7º/sec (8 pixels per frame).

Procedures

Thresholds were measured using a staircase procedure that tracked the 75% correct level of performance in our two-interval forced-choice task (see Lelkens & Koenderink, 1984). In this procedure, the log SNR value for a particular motion condition always changed following a trial, but by an amount that depended upon the subject's response. If the subject responded correctly on a particular trial, the log SNR was reduced by 1 step the next time that condition was presented. If the subject responded incorrectly, the log SNR for the subsequent trial was incremented by 3 steps. The initial step size for any motion condition was always set to a value of 0.4 log SNR and then declined by half with each response reversal until a constant minimum value of 0.05 was reached. Thresholds for motion detection in Experiment 2 were based on averages of the last 36 of 40 staircase reversals, when the log SNR step size was at its minimum value. In any experimental session, the particular combination of stimulus speed, field size, duration, and texture element size was held constant, and thresholds were measured for each of the five motion conditions in the standard suite.

Results

Log threshold SNRs for the standard suite of motion conditions, as a function of the speed of the two signal layers in the stimuli, are shown in Figures 3a–3c. Thresh-

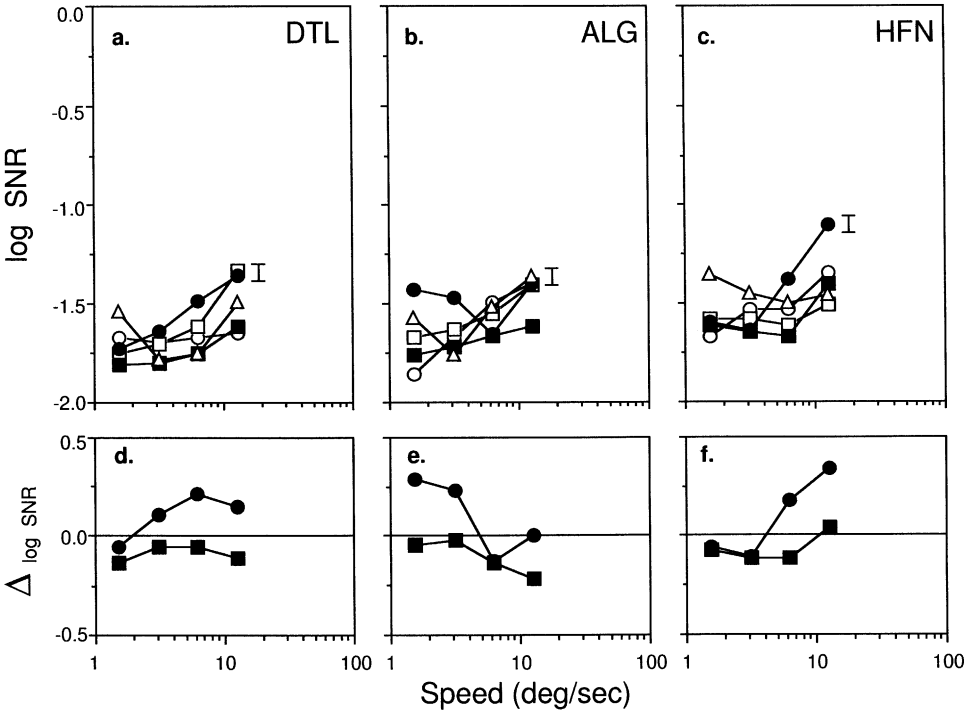

Figure 3. Signal-detection thresholds (log SNR) for unidirectional and bidirectional random-check textures as a function of the speed of the signal layers. Field size = 6.8° × 6.8°, duration = 2.67 sec, element size = 0.027°. Data for 3 subjects are shown. Top panels (a–c): log threshold SNRs for the standard suite of five signal-layer conditions: right only, left only, down only, left right, and down left. Consult legend in Figure 2 for identification of plotting symbols. The error bar to the immediate right of the data plot indicates ±1 standard error based on averaging standard error for each subject across all conditions in Experiment 2: ±0.029 (D.T.L.), ±0.029 (A.L.G.), ±0.030 (H.F.N.). Bottom panels (d–f): Data for the left-right (filled circles) and down-left (filled squares) compound-motion conditions, replotted as differences from the average of the unidirectional-motion threshold SNRs at each speed. Consult text for further details.

olds tended to remain relatively constant across the range of speeds studied, varying by no more than about 0.5 log SNR for any subject. Nearly all of the data across the 3 subjects tested fell in the range of −1.25 to −1.75 log SNR. The highest threshold SNR for each motion condition occurs at the highest velocity tested, 12.7°/sec, a finding that holds for all 3 subjects tested. Otherwise, there are no specific trends in the relationship between threshold log SNR and pattern speed that hold for all 3 subjects. It should be kept in mind that thresholds for the different speeds were obtained on different days, and the absence of consistent trends among the 3 subjects may be due to daily random variation in sensitivity to motion.

An important feature of the detection thresholds depicted in the lower panels of Figure 3 is the relationship between the thresholds for opponent and nonopponent transparent motion. Log threshold SNR for the nonopponent stimulus (filled squares) was generally less than the corresponding threshold for opponent motion (filled circles). Figures 3d–3f show the opponent and nonopponent transparent-motion thresholds in terms of Δlog SNR, the difference in threshold between the transparent-motion condition and the average for the three unidirectional motion conditions obtained in the same session: $\log \mathrm{SNR}_{\mathrm{T,directional}} - \log \mathrm{SNR}_{\mathrm{T,unidirectional}}$. We refer to Δlog SNR as relative thresholds. Positive values of Δlog SNR indicate that motion in the bidirectional condition was relatively less detectable than suggested by the average threshold SNR for the unidirectional conditions. The average differences in relative threshold across the four speeds examined in Experiment 2 were 0.19, 0.20, and 0.16 log SNR for D.T.L., A.L.G., and H.F.N., respectively. The results of a 4 speeds × 2 bidirectional motion types analysis of variance (ANOVA) with replications indicated significant differences in relative threshold between the opponent and perpendicular motion stimuli [$F(1,16) = 10.5, p < .01$] Neither the effects of speed [$F(3,16) = 0.134, p > .75$] nor the effects of the interaction of speed and motion type [$F(3,16) = 0.236$, $p > .75$] were significant.

The effects of stimulus duration on motion detection for random-check patterns moving at 12.7°/sec are shown in Figure 4. Log threshold SNR declines as stimulus duration increases, as seen in the Figure 4a–4c. The diagonal line of slope −0.5 drawn in Figure 4d provides a reference for gauging the rate of improvement in threshold

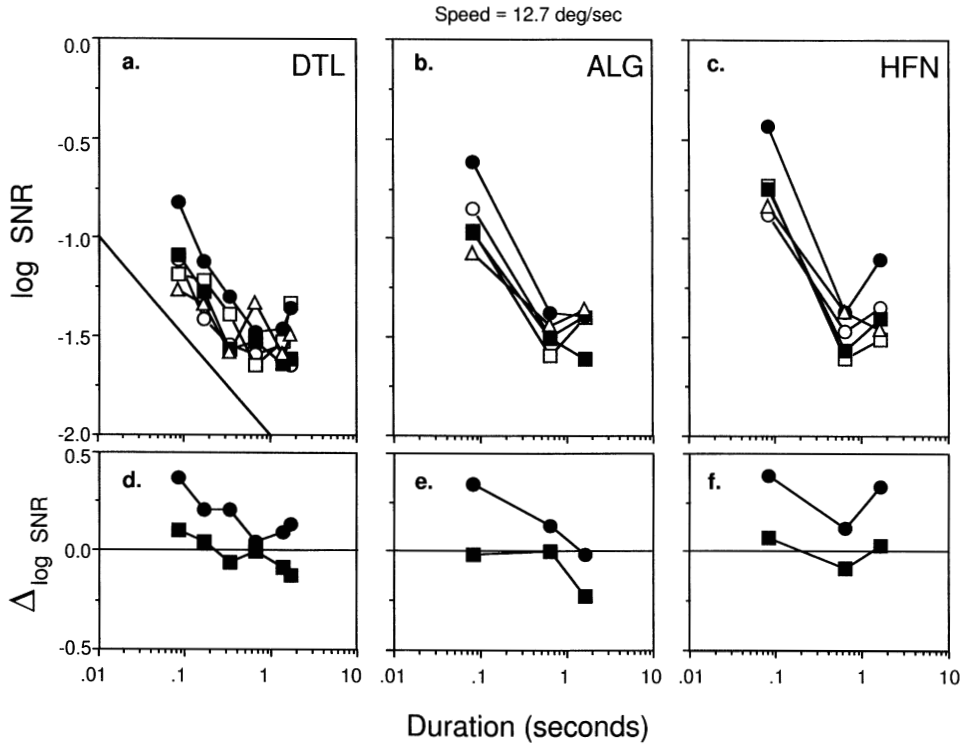

Figure 4. Signal-detection thresholds (log SNR) for unidirectional and bidirectional random-check textures as a function of the duration of the animation sequences. Speed = 12.7°/sec, field size = 6.8°, element size = 0.027°. Data for 3 subjects are shown. Consult Figure 3 and text for additional details.

SNR with duration. A slope of −0.5 is expected of an ideal motion detector, one based on cross-correlation of pixel luminances across space and time (Lappin & Bell, 1976; van Doorn & Koenderink, 1984). In the case of Subject D.T.L., for whom a complete set of thresholds was measured, log threshold SNR declines approximately as the square root (slope = −0.48) of duration for durations less than approximately 0.67 sec. Threshold does not continue to improve at this rate for longer durations and may actually asymptote to a final value by 1 sec.

Log relative thresholds for the two bidirectional motion stimuli are shown in Figures 4d–4f. Relative threshold did not vary with stimulus duration in any way that was consistent with subject. Perpendicular motion thresholds, however, were always lower than the corresponding opponent-motion thresholds. The average log differences in relative threshold were 0.20, 0.23, and 0.27 for Subjects D.T.L., A.L.G., and H.F.N., respectively. The results of a 3 (durations) × 2 (bidirectional motion types) ANOVA with replications revealed significant differences in relative threshold between opponent and perpendicular motion stimuli [$F(1,12) = 27.0, p < .001$]. The effects of duration were not significant [$F(2,12) = 1.58$, $p > .50$], although there was a significant interaction of duration and motion type [$F(2,12) = 7.31, p < .02$].

The 1.6°/sec data are shown in Figures 5a–5c. Threshold SNR declined as stimulus duration increased. The results for the bidirectional and unidirectional motion conditions closely paralleled one another. The threshold SNRs for the unidirectional motion conditions replicate previous results obtained by others in two important respects. First, the range of thresholds was similar to those reported previously by van Doorn and Koenderink (1982b), among others. Second, unidirectional thresholds in Figure 5 improved more rapidly with duration than did those measured at 12.7°/sec, a finding that is consistent with Fredericksen, Verstraten, and van de Grind's (1993,1994) results. The bottom panels in Figure 5 (d–f) indicate that the differences in log relative threshold SNR between opponent-motion and perpendicular-motion stimuli obtained at 12.7°/sec were also seen at the lower speed in 2 of 3 subjects. These differences were 0.14, 0.21, and 0.0 for D.T.L., A.L.G., and H.F.N., respectively. However, the results of a 3 (durations) × 2 (bidirectional motion types) ANOVA with replications did not show significant differences in relative threshold between opponent and perpendicular motion stimuli [$F(1,12) = 3.42, p > .10$]. Nor were the effects of duration [$F(2,12) = 0.115, p > .75$] or the interaction of duration and motion type [$F(2,12) = 0.096, p > .75$] significant.

The effect of stimulus area on the detectability of motion for the suite of motion conditions is shown in Figures 6a–6c. The thresholds for detecting motion in unidirectional stimuli improved as field size was increased. The diagonal line of slope −0.5 drawn in Figure 6d indicates that the improvement approximately followed a

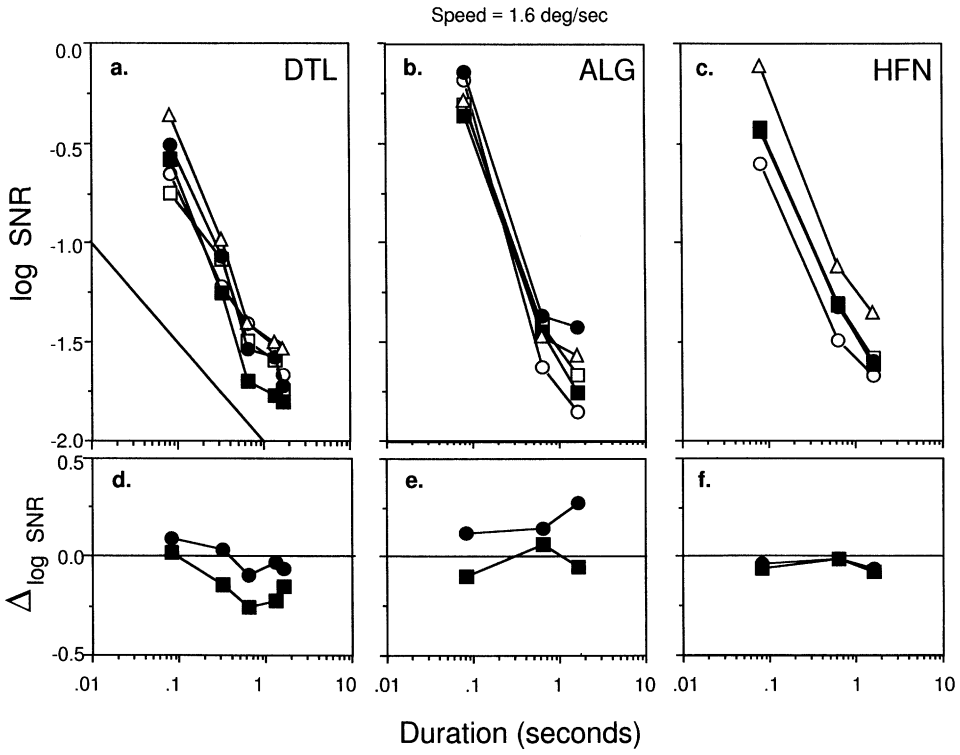

Figure 5. Signal-detection thresholds (log SNR) for unidirectional and bidirectional random-check textures as a function of the duration of the animation sequences. Speed = 1.6°/sec, field size = 6.8° × 6.8°, element size = 0.027°. Data for 3 subjects are shown. Consult Figure 3 and text for additional details.

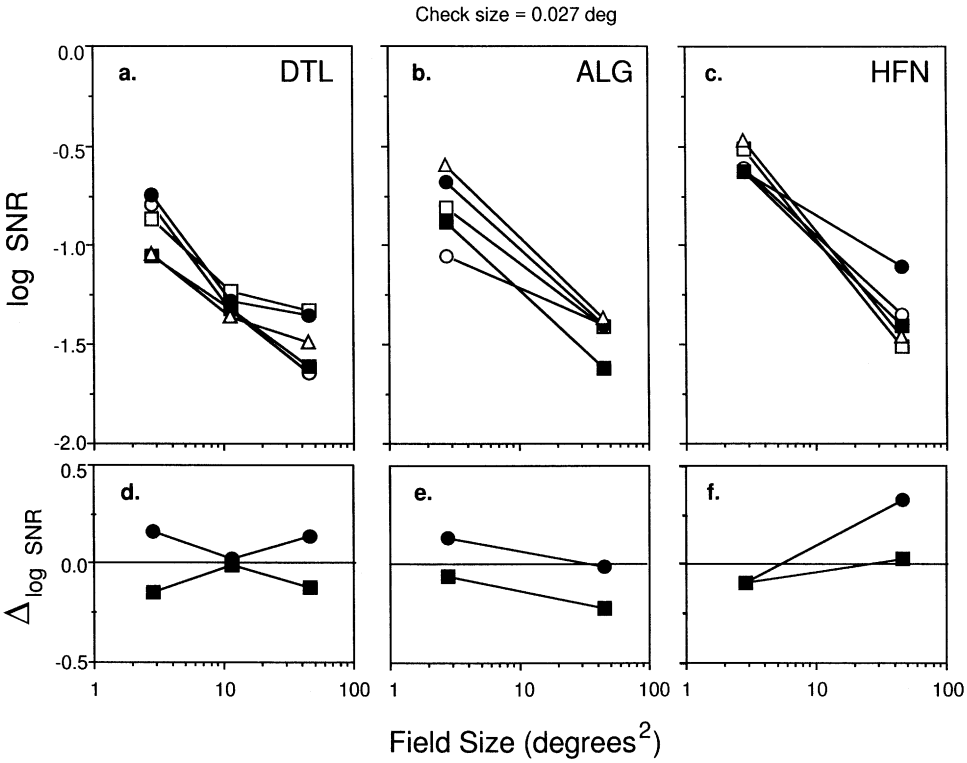

Figure 6. Signal-detection thresholds (log SNR) for unidirectional and bidirectional random-check textures as a function of the stimulus area. Speed = 12.7°/sec; duration = 2.67 sec; element size = 0.027°. Data for 3 subjects are shown. Consult Figure 3 and text for additional details.

square-root law. Thresholds for motion detection in bidirectional stimuli also showed improvement as field size was increased, and these improvements paralleled those obtained for unidirectional motion. Detection threshold was always higher for the opponent stimulus than for the nonopponent stimulus. Thresholds for the two stimulus conditions differed from one another, on the average, by 0.20, 0.21, and 0.15 log units for D.T.L., A.L.G., and H.F.N., respectively. The results of a 2 (sizes) $\times$ 2 (compound motion types) ANOVA with replications revealed a significant difference in relative threshold between opponent and nonopponent motion stimuli [$F(1,8) = 8.11$, $p < .05$]. Neither the effects of stimulus area size [$F(1,8) = 0.299$, $p > .50$] nor the interaction of stimulus area and motion type [$F(1,8) = 0.372$, $p > .50$] were significant.

Figure 7 shows results for a single subject, in which the stimuli were composed of checks that were 8 pixels on a side. Comparison of these data with those in Figures 6a and 6d reveals little effect of check size under the conditions in which thresholds were obtained in this experiment: thresholds obtained with the coarser texture patterns were generally lowest for the perpendicular motion condition, middling for the unidirectional conditions, and highest for the opponent motion conditions. Moreover, this pattern of results was sustained for patterns as large as 184.4 deg^2.

In summary, the results of Experiment 2 suggest that the pattern of results obtained in Experiment 1 generalizes to large variations in stimulus velocity (8::1), field size (16::1), duration (20::1), and check size (8::1). In all but one condition examined, the motion-detection threshold for the nonopponent-motion condition was lower than the threshold for the opponent-motion condition. The effects, however, were modest, amounting to an average difference of approximately 0.19 log units (a factor of 1.6:1). We also find that, *on the average*, detection thresholds for unidirectional motion tended to fall between those obtained for the two bidirectional motion conditions. This trend in the data is clearly evident in the correlation plot shown in Figure 8, which replots all of the data in Experiment 2 in a way that permits comparison of the relative thresholds for perpendicular motion and opponent motion. Data for all 3 subjects are combined in the plot. If subjects were equally sensitive to opponent and perpendicular motion conditions, we would expect the data to be distributed along the diagonal line shown in the figure. Instead, the data tend to cluster below the diagonal, indicating a lower relative threshold for perpendicular motion than for opponent motion. The asterisk in Figure 8 corresponds to the average results for all conditions tested in Experiment 2. Although there is some variability in results across the various conditions and subjects tested, the asterisk shows that, on average, subjects were more sensitive to perpendicular motion than to unidirectional motion, and they were less sensitive to opponent motion than to unidirectional motion.

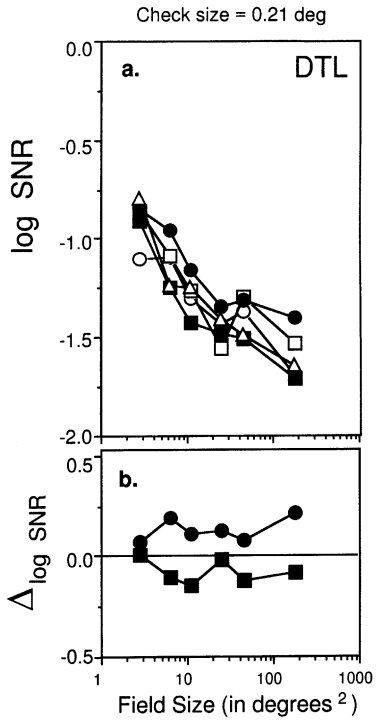

Figure 7. D.T.L.'s signal-detection thresholds (log SNR) for unidirectional and bidirectional random-check textures as a function of the stimulus area. Speed = 12.7°/sec; duration = 2.67 sec; element size = 0.21°. Consult Figure 3 and text for additional details.

EXPERIMENT 3

The results of the preceding two experiments provide compelling evidence in favor of direction-specific inhibitory or suppressive interactions in the perception of transparent motion. However, we felt that Experiment 3 was warranted by one important issue concerning these effects that our motion-detection procedures did not specifically address. As pointed out in the General Methods section of this paper, the interpretation of the differences in the detectability of motion in the uni- and bidirectional conditions depends to some extent on a model of signal detectability that accounts for the effects of probability summation. We wanted to run a control experiment in which we could assay for the suppression of one motion signal by another without the confounding effects of probability summation. In Experiment 3, subjects were asked to identify one of the directions of motion—left, right, up, or down—in a stimulus. In some trials, the motion was unidirectional and the target direction corresponded to the stimulus direction; in other trials, the motion was bidirectional and the target direction corresponded to one or the other component in the stimulus. The identification task that we designed is an appropriate control, because subjects could base their forced-choice judgments solely on a motion signal in a single direction in both uni- and bidirectional conditions. Also, we designed the task so that it, like the detection task, would be a forced-choice procedure in which subjects could choose be-

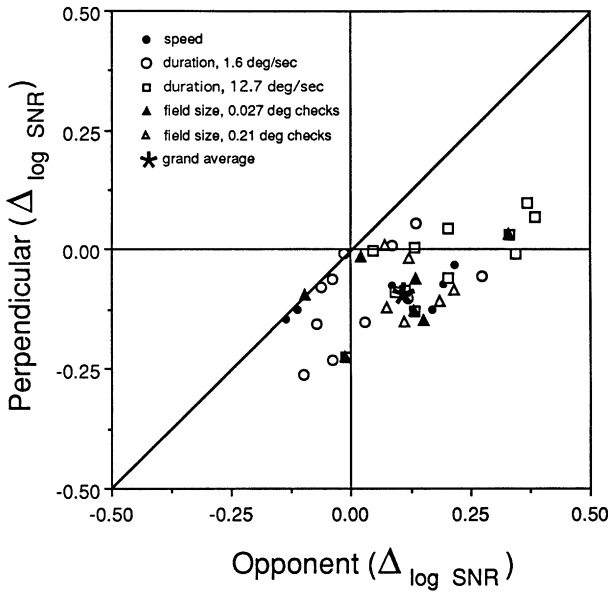

Figure 8. Correlation plot of all log relative thresholds, Δlog SNR, for perpendicular- versus opponent-motion, obtained in Experiment 2. Relative thresholds obtained as a function of speed, duration, field size, etc., are plotted with different symbols, as indicated in the figure legend. The mean relative thresholds for opponent- and perpendicular-motion are indicated by the asterisk.

tween two response alternatives. However, rather than choosing between signal-plus-noise and noise-only intervals as in Experiments 1 and 2, subjects in Experiment 3 had to choose between a target direction of motion and a distractor direction in a single stimulus interval per trial.

Method

The dynamic random-check patterns used in Experiment 3 were of the same size, duration, and speed, and were presented within the same temporal contrast window as those presented in Experiment 1. Identification thresholds were measured for all 16 possible combinations of the two signal layers and four directions of motion: left, right, up, and down. Thresholds were measured using the following forced-choice paradigm. Following the presentation of a stimulus, the subjects were presented with a choice of two arrows, pointing in different directions, drawn on the computer monitor. One of the arrows, the target, pointed in a direction that matched a direction of motion in the animation sequence. The other arrow, the distractor, pointed in a direction that did not correspond to any motion component in the stimulus. The subjects used the computer mouse to select one of the arrows. Arrows corresponding to the four directions of motion in the experiment served as distractors with equal mean frequency. Thresholds were measured using the same 3-up/1-down staircase procedure used in Experiment 2. The 16 individual staircases were all interleaved, and trials on each staircase were presented in block-random fashion. The reported thresholds are all based on 36 reversals collected over the course of two experimental sessions.[5]

Results

Average log thresholds for the correct identification of motion in our displays are shown in Figure 9. The data from the 16 conditions tested in Experiment 3 have been collapsed into three categories labeled "0," "90," and "180" in the figure. The category names refer to the differences, in degrees, between the directions of motion of the two signal layers and correspond to unidirectional, perpendicular, and opponent motion conditions, respectively. The error bars in the figures correspond to 1 standard error of the mean. Motion identification thresholds for all subjects were similar to the corresponding detection thresholds measured in Experiments 1 and 2, although A.L.G.'s identification thresholds for opponent motion are, on average, 0.4 log SNR higher than the corresponding detection thresholds.

It is clear from inspection of Figure 9 that there were no differences in the subjects' performance on the 0° and 90° stimuli, and that the subjects performed less well on the 180° stimuli than they did on stimuli in either of the other two categories. The results of an ANOVA with planned comparisons among the 0°, 90°, and 180° conditions collapsed across subjects confirmed this observation ($p < .05$). The differences between 0° or 90° and 180° identification thresholds averaged 0.3 log units, or about a factor of 2, across the 3 subjects. Thus, the results of Experiment 3, like those of the previous two experiments, are consistent with the view that inhibitory or suppressive interactions in bidirectional motion are direction-specific: they affect detection and identification to a greater extent when the signal layers move in opposite directions than when they move at right angles to one another. Furthermore, the results of Experiment 3 indicate little if any effect of suppression in the perpendicular bidirectional motion condition. The modest improvement in sensitivity for perpendicular motion (as compared with unidirectional motion) observed in Experiment 1 apparently reflects signal summation processes alone rather than a combination of competing summation and suppressive processes.

DISCUSSION

In the introduction to this article, we pointed out that moving objects stimulate diverse populations of motion sensors, which are tuned to different spatial frequencies and orientations, and temporal frequencies of motion. A fundamental unsolved problem in vision is how information from all these various sources is integrated into a unified representation of moving objects that is approximately veridical. Local inhibitory interactions among sensors tuned to different directions of motion have been proposed as a process contributing to image segmentation and unification.

A number of authors have previously pointed out that the computation of motion transparency is based in part on global properties other than the velocities of the transparent objects—for example, spatial frequency, contrast, color, and binocular disparity (e.g., Adelson & Movshon, 1982; Kooi, DeValois, Switkes, & Grosof, 1992; Krauskopf & Farell, 1990; Krauskopf, Wu, & Farell, 1996;

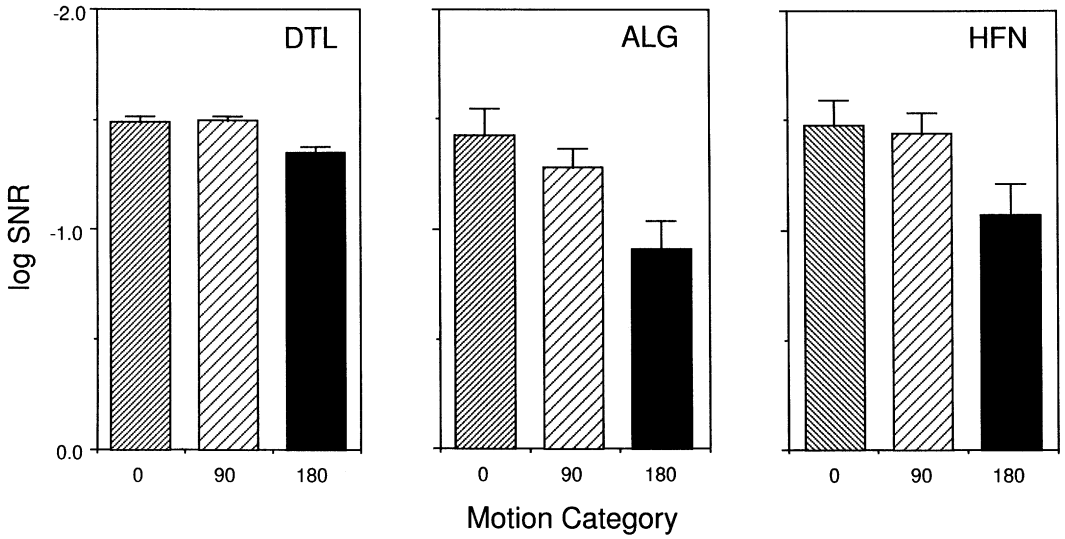

Figure 9. Log threshold SNRs for the identification of direction of motion in dynamic random-check patterns. Results for 3 subjects are shown. The data from the 16 conditions tested in Experiment 3 have been collapsed into three categories: "0," "90," and "180." The bars in each panel represent the logarithm of the average threshold SNR for identification of leftward, rightward, upward, and downward motion in each of three stimulus configurations. Category 0, both signal layers moved in same direction; category 90, signal layers moved in directions perpendicular to one another; category 180, signal layers moved in opposite directions. Error bars: 1 standard error of the mean threshold for leftward, rightward, upward, and downward identifications.

Movshon, Adelson, Gizzi, & Newsome, 1985; Qian et al., 1994a; Stone, Watson, & Mulligan, 1990). These authors have shown that differences among moving transparent objects along one or more of these dimensions can provide a simple means for grouping the motion signals appropriately. Other investigators have observed that higher level perceptual processes, such as those that may be responsible for perception of the surface properties of objects, can facilitate the segmentation of overlapping moving patterns into their constituent parts (e.g., Lindsey & Todd, 1996; Stoner & Albright, 1993, 1996; Stoner, Albright, & Ramachandran, 1990; Trueswell & Hayhoe, 1993). While the visual system may be capable of using these cues in the analysis of moving stimuli, it is also clear that these cues are not necessary for motion transparency to be perceived. The individual layers in our dynamic random-check stimuli are indistinguishable from one another on the basis of any of these potential cues, although motion transparency is readily perceived in these stimuli.

The results of the present study also make clear three additional points concerning the processes occurring in the perception of motion transparency in random-check patterns. First, they indicate that the statistical efficiency with which visual processes extract motion from the bidirectional motion stimuli may depend to some extent upon the relative directions of motion of the two signal layers. The results of the first two experiments clearly show across a broad range of stimulus parameters that motion in opponent-motion stimuli is more difficult to detect when it is embedded in noise than is motion in either unidirectional- or perpendicular-motion conditions. Second, the range of differences among the motion conditions tested is very modest. The largest differences in performance are on the order of a factor of 2. Third, the results of the present study also indicate that the differences in sensitivity are not determined solely by whether the signal layers move in the same or different directions. Although sensitivity was generally lowest for opponent-motion stimuli, sensitivity was generally highest for the perpendicular-motion stimuli.

Several other studies have previously examined the detectability of opponent motion. Van Doorn and Koenderink (1982a, 1982b) measured SNRs for stimuli consisting of dynamic random-check noise superimposed on interdigitated rows of leftward- or rightward-moving random checks. Motion of the individual rows was not resolved when the rows were a single pixel in height, and subjects perceived left-right transparent motion, provided the patterns were viewed at suprathreshold values of SNR. Lappin and Kottas (1981) compared the coherence-detection thresholds in sparse random-dot lattices moving in one direction as functions of the direction of a background lattice moving either perpendicular to or opposite the first pattern. Mather and Moulden (1983) compared luminance thresholds for detecting motion in sparse random-dot displays consisting of either a single lattice in unidirectional motion or two lattices in opponent motion. The results of all these studies, obtained across a variety of different techniques, are in agreement with those of the present study in finding that motion in the opponent condition is less detectable and less identifiable than motion in a unidirectional comparison condition. Therefore, the psychophysical evidence for inhib-

itory or suppressive interactions in opponent-motion stimuli seems firmly established.

The extent to which similar interactions occur in response to bidirectional motion when the components are moving at right angles to one another is less well established. Although we do not find evidence for inhibitory interactions in these conditions, Snowden (1989, 1990), using superimposed sparse random-dot arrays, found that the minimum and maximum detectable displacements (d_{min} and d_{max}) of a horizontally moving array were adversely affected by the addition of a vertically moving array. These interactions were dependent upon the relative speeds of the component random-dot arrays (Snowden, 1990). Verstraten, Fredericksen, van Wezel, Boulton, and van de Grind (1996) compared SNRs for direction of motion discriminations of a foreground random-check texture that was interdigitated with background textures that were either stationary or moving at right angles to the test texture. Threshold SNRs were elevated by the presence of the moving background textures, although Verstraten et al., like Snowden (1990) found that these effects were speed-dependent.

In comparing these results with those of the present study, it is important to recognize that both Snowden and Verstraten et al. employed high-contrast background stimuli. In the present series of experiments, however, the SNR thresholds for bidirectional motion were always measured under conditions in which background contrast was, logically, near threshold, since foreground and background contrasts were always equal (see Equation 1). Therefore, the results of the present study should be regarded as an analysis of interactions among weak motion signals, and, under these conditions, our experiments reveal little interaction in perpendicular motion. Under conditions in which signal levels are large, a mutual partial cancellation of motion signal is found in both opponent- and perpendicular-motion conditions. In fact, when the motion signals produced by bidirectional stimuli are balanced locally (Qian et al., 1994a; see also Mace & Shaw, 1974), the motion signals that remain following cancellation are too weak to sustain the perception of motion. The dependence of signal cancellation on the contrasts of the component stimuli has been previously pointed out by Stromeyer et al. (1984) in connection with the detectability of contrast increments in cosine gratings moving in opposite directions. At very low grating contrasts, the detection of contrast increments appears to be governed by independent, direction-specific motion mechanisms (see also Watson, Thompson, Murphy, & Nachmias, 1980), and no influence of inhibitory interactions is seen. At high grating contrasts, detection of contrast increments was influenced by what Stromeyer et al. concluded were inhibitory effects between mechanisms tuned to opposite directions of motion.

How might visual processes be capable of segregating overlapping transparent motions? A number of years ago, Lappin and Bell (1976) suggested that the computation of global motion might involve processes closely paralleling cross-correlation. This proposal was based in part on their finding that the detection of motion in two-frame animation sequences involving sparse arrays of random dots improves as the square root of the number of dots. Later, van Doorn and her colleagues (e.g., van de Grind, Koenderink, & van Doorn, 1987; van Doorn & Koenderink, 1982a, 1982b, 1984) used a similar conceptual approach as a guide in their extensive studies of the detectability of motion in the presence of noise. While cross-correlation provides a computational basis for the perceptual segregation of bidirectional random-check stimuli into their moving constituents, it does not require the existence of suppressive or inhibitory interactions across sensors tuned to different directions of motion. The segregation of transparent motion by cross-correlation is based on the idea that the random-check patterns provide two statistically independent sources of motion information. Cross-correlation readily identifies these two independent sources by linear integration across space and time of the local correlations in each of a number of different hypothesized magnitudes and directions of random-check displacements. The estimates of source velocities derived from cross-correlation will correspond to those velocities for which correlation values are statistically significantly different from the background values at other velocities.

More recently, a number of alternative models based on spatiotemporally oriented linear filter pairs arranged in quadrature spatiotemporal phase have been proposed for the computation of global motion. The sums of squared outputs of these filters provide estimates of motion energy (Adelson & Bergen, 1985) within a particular spatiotemporal frequency band. Different populations of sensors provide measurements in different bands. Unlike the cross-correlation model discussed above, models based on the measurement of spatiotemporally oriented energy cannot solve the transparency problem by linear integration, and other strategies must be devised. This point is illustrated in Figure 10, which shows idealized Fourier amplitude spectra for unidirectional-, opponent-, and perpendicular-motion random-check patterns. The axes in the figure indicate 2-D spatial frequencies and temporal frequency (f_x, f_y, and f_t, respectively). As shown in the figure, the energies of moving patterns are confined to planes within 3-D Fourier space. For the idealized case depicted in Figure 10, an average of many stimuli is assumed and the individual checks of the moving patterns are assumed to be vanishingly small. Therefore, the energy is uniformly distributed on the plane.

The idealized power spectrum for unidirectional, rightward motion is shown in Figure 10a. This stimulus will elicit responses in sensors with spatiotemporal tunings that lie on or near planes; those sensors tuned to other frequency bands will respond feebly if at all. The exact pattern of responses among motion sensors will depend upon the orientation of the plane in Fourier space, which in turn will depend on the velocity of the moving texture. The pattern of responses among motion sensors elicited

Figure 10. Idealized 3-D Fourier power spectra for random-check patterns consisting of (a) two signal layers moving leftward at 1°/sec, (b) 1°/sec leftward- plus rightward-moving signal layers, and (c) 1°/sec leftward- and downward-moving signal layers. In each case, a noise-layer contrast of zero is assumed. F_x and F_y are in cycles/deg; F_t is in hertz.

by this planar distribution of energy can be used to estimate the velocity of the moving texture pattern, as Heeger (1987), among others, has demonstrated. However, two fundamental problems arise when this sort of approach is attempted for the two transparent motion cases depicted in Figures 10b and 10c. First, the motion-energy measurements now arise from two moving patterns rather than from one pattern, and a way must be found to appropriately group the responses into those arising from each of the moving layers before computation of their separate velocities can proceed. Second, algorithms that might base the segregation of these responses on maxima in the local measurements of motion energy will not work reliably, since these maxima occur, on average, in spatiotemporal frequency bands centered on the intersections of the planes associated with each of the moving patterns. In the case of opponent motion depicted in Figure 10b, this is not a problem, since these maxima lie in the f_x, f_y, plane at right angles to the dominant directions of motion and are consistent with zero-velocity patterns. In the case of perpendicular motion depicted in Figure 10c, however, the maximum lies on a diagonal line in the down-left direction corresponding to the vector average of the two pattern velocities.

A potential solution to these two problems has been proposed by Qian et al. (1994a, 1994b). According to their proposal, sensors with similar frequency tuning but different directional tuning characteristics are mutually inhibitory or suppressive, while sensors with similar directional tuning characteristics but different frequency tuning will facilitate one another. These interactions occur only among signals that arise approximately coincidentally in space and time, and they lead to a single resultant estimate of velocity at each location in the stimulus field. According to this view, (1) motion transparency in random-check patterns occurs because much of the motion energy associated with each of the two layers is uncorrelated, even though the two layers have similar space- and time-averaged motion energies; (2) in the case of perpendicular motion, much of the motion energy measured in the diagonal direction is cancelled because it occurs coincidentally with motion energy in the direction of either or both component patterns; and (3) spatiotemporally local interactions lead predominantly to interdigitated regions in the representation of the visual field in which the resultant direction of encoded motion is either leftward or downward.

The proposal by Qian et al. (1994a, 1994b) seems to fit a number of aspects of our empirical results when we compare unidirectional- and opponent-motion conditions. On the other hand, the fit between empirical result and theory is less perfect when we consider perpendicular motion. A reasonable interpretation of our results is that the difference in performance between opponent and perpendicular motion conditions reflects a direction-specific difference in the effectiveness of suppression in the interaction stage, as proposed by Qian et al. According to this view, subjects' greater sensitivity to perpendicular motion may be due to incomplete cancellation of motion in the down-left diagonal direction, as compared with the more effective cancellation of spatiotemporally coincident motion in the opponent-motion conditions. A potentially serious problem with this view, however, is that it suggests that subjects should see diagonal motion when viewing perpendicular-motion random-check patterns, since it is that direction that contains the greatest fraction of motion energy, as shown in Figure 10c. None of the subjects in our study ever reported seeing diagonal motion during any of the experimental sessions. In fact, van Doorn and Koenderink (1982a) have reported that transparent motion is seen in random-check layers moving at equal speeds, even when the layers are sepa-

rated by as little as 30°. It is not clear how locally suppressive or inhibitory interactions, by themselves, can account for this degree of selectivity.

In conclusion, we have used a signal-detection paradigm to reveal direction-specific inhibitory or suppressive interactions among near-threshold signals occurring among motion sensors responding to transparent motion stimuli. These interactions occur across a wide variation in stimulus speed, field size, duration, and texture element size. Many, though not all, features of our data are consistent with a recent proposal that bidirectional motion stimuli may be segregated into their component velocities by spatiotemporally local interactions in which sensors tuned to similar directions of motion are mutually facilitating, while those tuned to different directions of motion are mutually inhibitory or suppressive.

REFERENCES

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

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

Andersen, G. J. (1989). Perception of three-dimensional structure from optic flow without locally smooth velocity. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 363-371.

Andersen, G. J., & Wuestefeld, A. D. (1993). Detection of three-dimensional surfaces from optic flow: The effects of noise. *Perception & Psychophysics*, **54**, 321-333.

Balch, W., & Shaw, R. E. (1978). The role of perceptual organization in the depth perception of kinetic lattice displays. *Perception & Psychophysics*, **23**, 145-152.

Barlow, H. B., & Levick, W. R. (1965). The mechanism of directionally selective units in rabbit's retina. *Journal of Physiology*, **178**, 477-504.

Clarke, P. G. H. (1977). Subjective standstill caused by the interaction of moving patterns. *Vision Research*, **17**, 1243.

Fredericksen, R. E., Verstraten, F. A. J., & van de Grind, W. A. (1993). Spatio-temporal characteristics of human motion perception. *Vision Research*, **33**, 1193-1205.

Fredericksen, R. E., Verstraten, F. A. J., & van de Grind, W. A. (1994). Temporal integration of random dot apparent motion information in human central vision. *Vision Research*, **34**, 461-476.

Gibson, E. J., Gibson, J. J., Smith, O. W., & Flock, H. (1959). Motion parallax as a determinant of perceived depth. *Journal of Experimental Psychology*, **58**, 40-51.

Graham, N. (1989). *Visual pattern analyzers*. New York: Oxford University Press.

Green, D. M., & Swets, J. A. (1966). *Signal detection and psychophysics*. New York:Wiley.

Grzywacz, N. M., & Yuille, A. L. (1990). A model for the estimate of local image velocity by cells in the visual cortex. *Proceedings of the Royal Society of London: Series A*, **239**, 129-161.

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

Hildreth, E. C. (1984). Computations underlying the measurement of visual motion. *Artificial Intelligence*, **23**, 309-355.

Hiris, E., & Blake, R. (1996). Direction repulsion in motion transparency. *Visual Neuroscience*, **13**, 187-197.

Horn, B. K. P., & Schunk, B. G. (1981). Determining optical flow. *Artificial Intelligence*, **17**, 185-203.

Jasinschi, R. , Rosenfeld, A., & Sumi, K. (1992). Perceptual motion transparency: The role of geometrical information. *Journal of the Optical Society of America A*, **9**, 1865-1879.

Kim, J., & Wilson, H. R. (1992). Dependence of plaid motion coherence on component grating directions. *Vision Research*, **33**, 2479-2489.

Koenderink, J. J., van Doorn, A. J., & van de Grind, W. A. (1985). Spatial and temporal parameters of motion detection in the peripheral visual field. *Journal of the Optical Society of America A*, **2**, 252-259.

Kooi, F., DeValois, K. K., Switkes, E., & Grosof, X. (1992). Higher order factors influencing the perception of sliding and coherence of a plaid. *Perception*, **21**, 583-598.

Krauskopf, J., & Farell, B. (1990). Influence of color on the perception of coherent motion. *Nature*, **348**, 328-331.

Krauskopf, J., Wu, H.-J., & Farell, B. (1996). Coherence, cardinal directions and higher-order mechanism. *Vision Research*, **36**, 1235-1246.

Lappin, J. S., & Bell, H. H. (1976). The detection of coherence in moving random-dot patterns. *Vision Researc*h, **48**, 161-168.

Lappin, J. S., & Kottas, B. L. (1981). The perceptual coherence of moving visual patterns. *Acta Psychologica*, **48**, 163-174.

Lelkens, A. M. M., & Koenderink, J. J. (1984). Illusory motion in visual displays. *Vision Research*, **24**, 1083-1090.

Lindsey, D. T., & Todd, J. T. (1996). On the relative contributions of motion energy and transparency to the perception of moving plaids. *Vision Research*, **36**, 207-222.

Mace, W. M., & Shaw, R. E. (1974). Simple kinetic information for transparent depth. *Perception & Psychophysics*, **15**, 201-209.

Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.

Marshak, W., & Sekuler, W. (1979). Mutual repulsion between moving visual targets. *Science*, **205**, 1399-1401.

Mather, G. (1980). The movement aftereffect and a distribution-shift model for coding the direction of visual movement. *Perception*, **9**, 379-392.

Mather, G., & Moulden, B. (1983). Thresholds for movement direction: Two directions are less detectable than one. *Quarterly Journal of Experimental Psychology*, **35A**, 513-518.

Movshon, J. A., Adelson, E. H., Gizzi, M. S., & Newsome, W. T. (1985). The analysis of moving visual patterns. In C. Chagas, R. Gattass, & C. Gross (Eds.), *Pattern recognition mechanisms* (Pontificae Academicae Scientiarum Scripta Varia, Vol. 54, pp. 117-151). Rome: Vatican Press.

Nowlan, S. J., & Sejnowski, T. J. (1995). Filter selection model for motion segmentation and velocity integration. *Journal of the Optical Society of America A*, **11**, 3177-3200.

Pelli, D. G. (1985). Uncertainty explains many aspects of visual contrast detection and discrimination. *Journal of the Optical Society of America A*, **2**, 1508-1532.

Poggio, T., Yang, W., & Torre, V. (1988). Optical flow: Computational properties and networks, biological and analog. In R. Durban, C. Miall, & G. Mitcheson (Eds.), *The computing neuron* (pp. 355-370). Wokingham, U.K.: Addison-Wesley.

Qian, N., Andersen, R. A., & Adelson, E. H. (1994a). Transparent motion perception as detection of unbalanced motion signals: I. Psychophysics. *Journal of Neuroscience*, **14**, 7357-7366.

Qian, N., Andersen, R. A., & Adelson, E. H. (1994b). Transparent motion perception as detection of unbalanced motion signals: III. Modeling. *Journal of Neuroscience*, **14**, 7381-7392.

Reichardt, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 303-317). Cambridge, MA: MIT Press.

Snowden, R. J. (1989). Motions in orthogonal directions are mutually suppressive. *Journal of the Optical Society of America A*, **6**, 1096-1101.

Snowden, R. J. (1990). Suppressive interactions between moving patterns: Role of velocity. *Perception & Psychophysics*, **47**, 74-78.

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

Stoner, G. R., & Albright, T. D. (1993). Image segmentation cues in motion processing: Implications for modularity in vision. *Journal of Cognitive Neuroscience*, **5**, 129-149.

Stoner, G. R., & Albright, T. D. (1996). The interpretation of visual

motion: Evidence for surface segmentation mechanisms. *Vision Research*, **36**, 1291-1310.
Stoner, G. R., Albright, T. D., & Ramachandran, V. S. (1990). Transparency and coherence in human motion perception. *Nature*, **344**, 153-155.
Stromeyer, C. F., III, Kronauer, R. E., Madsen, J. C., & Klein, S. A. (1984). Opponent-movement mechanisms in human vision. *Journal of the Optical Society of America A*, **1**, 876-884.
Trueswell, J. C., & Hayhoe, M. M. (1993). Surface segmentation mechanisms and motion perception. *Vision Research*, **33**, 313-328.
van de Grind, W. A., Koenderink, J. J., & van Doorn, A. J. (1986). The distribution of human motion detector properties in the monocular visual field. *Vision Research*, **26**, 797-810.
van de Grind, W. A., Koenderink, J. J., & van Doorn, A. J. (1987). Influence of contrast on foveal and peripheral detection of coherent motion in moving random-dot patterns. *Journal of the Optical Society of America A*, **4**, 1643-1652.
van de Grind, W. A., van Doorn, A. J., & Koenderink, J. J. (1983). Detection of coherent movement in peripherally viewed random-dot patterns. *Journal of the Optical Society of America*, **73**, 1674-1683.
van Doorn, A. J., & Koenderink, J. J. (1982a). Spatial properties of the visual detectability of moving spatial white noise. *Experimental Brain Research*, **45**, 189-195.
van Doorn, A. J., & Koenderink, J. J. (1982b). Temporal properties of the visual detectability of moving spatial white noise. *Experimental Brain Research*, **45**, 179-188.
van Doorn, A. J., & Koenderink, J. J. (1984). Spatiotemporal integration in the detection of coherent motion. *Vision Research*, **24**, 47-53.
van Doorn, A. J., Koenderink, J. J., & van de Grind, W. A. (1984). Limits in spatio-temporal correlation and the perception of visual movement. In A. J. van Doorn, W. A. van de Grind, & J. J. Koenderink (Eds.), *Limits in perception: Essays in honour of Maartin A. Bauman* (pp. 203-234). Utrecht: VNU Science Press.
van Doorn, A. J., Koenderink, J. J., & van de Grind, W. A. (1985). Perception of movement and correlation in stroboscopically presented noise patterns. *Perception*, **14**, 209-224.
van Santen, J. P. H., & Sperling, G. (1984). Temporal covariance model of human motion perception. *Journal of the Optical Society of America A*, **1**, 451-473.
van Santen, J. P. H., & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, **2**, 300-321.
Verstraten, F. A. J., Fredricksen, R. E., & van de Grind, W. A. (1994). Movement aftereffect of bi-vectorial transparent motion. *Vision Research*, **34**, 349-358.
Verstraten, F. A. J., Fredericksen, R. E., van Wezel, R. J. A., Boulton, J. C., & van de Grind, W. A. (1996). Directional motion sensitivity under transparent motion conditions. *Vision Research*, **36**, 2333-2336.
Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, **45**, 205-217.
Watson, A. B., & Ahumada, A. J. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A*, **2**, 322-342.
Watson, A. B., Thompson, P. G., Murphy, B. J., & Nachmias, J. (1980). Summation and discrimination of gratings moving in opposite directions. *Vision Research*, **20**, 341-347.
Wohlgemuth, A. (1911). On the aftereffect of seen movement. *British Journal of Psychology Monographs* (Suppl. No. 1).

NOTES

1. Fredericksen, Verstraten, and van de Grind (1993) have previously referred to the method of varying SNR by varying the ratio of signal and noise luminance contrast as "LSNR," to distinguish it from another technique in which signal-to-noise ratio is determined by the numbers of elements in a dynamic display that move coherently relative to the number that jump randomly as the animation sequence progresses.

2. The specification of bidirectional SNR given in Equation 5 assumes that motion in one direction does not serve as noise for motion sensors that mediate the detection of motion in the other direction. The degree to which this assumption holds true will depend upon the difference in directions of motion of the components as well as upon the directional passbands of the motion sensors. An alternative specification of bidirectional SNR,

$$\mathrm{SNR} = \frac{q^2\mathrm{VAR}\,(\mathbf{P}_{1\text{ or }2})}{r^2\mathrm{VAR}\,(\mathbf{P}_3) + q^2\mathrm{VAR}\,(\mathbf{P}_{2\text{ or }1})} = \frac{q^2}{r^2 + q^2},$$

is based solely on luminance variance at each point in the stimulus field, and implicitly assumes that motion in one direction serves as noise for motion sensors responding to motion in the other direction. With few exceptions, performance was measured under conditions where $r^2 \gg q^2$ and both Equation 5 and the preceding equation give nearly identical values for SNR.

3. The upper-bound estimate is based on what Graham (1989) refers to as an "intermixed-summation" experiment and a decision rule based on the summation of sensor channel outputs. Noise is assumed to be gaussian distributed. Furthermore, it is assumed that subjects do not know the directions in which motion will appear in any given 2-IFC trial, and will therefore always monitor each of three motion sensor channels, each tuned to one of the three principle directions of motion signal in a session: left, right, and down (see Method sections in Experiments 1 and 2 for details). The prediction of $d'_{\mathrm{bi}} = d'_{1,\mathrm{uni}} + d'_{2,\mathrm{uni}}$ differs from the often-cited vector magnitude rule,

$$d' = \left[\sum_{i=1,n} \left(d'_i\right)^2\right]^{\frac{1}{2}},$$

because the latter is based on the assumption that the subject always knows which sensor channels to monitor in any given trial. Facilitation is greater in the presence of signal uncertainty than when the subject knows which channels to monitor. In the former situation, the sum of variances in the three channels is constant, regardless of stimulus condition, because all three channels must be monitored on every trial due to stimulus uncertainty. When the signal condition is known, the greater channel variance for the bi- versus unidirectional-motion conditions partially offsets the benefits of increased signal level as the number of independent signal sources increases. We wish to emphasize the fact that the expected effects of facilitation also depend upon the choice of a decision rule. Changing the rule from channel-summation to maximum-channel response leads to a significant reduction in predicted effects of probability summation (see Graham, 1989; Pelli, 1985).

4. The signal values depicted in Figure 2 are relative to the normalization $S^2 + N^2 = 1$, where S^2 and N^2 are the signal and noise variances, respectively. Since $\mathrm{SNR} = S^2/N^2$, the signal values and SNR are related by the following expression: $S^2 = \mathrm{SNR}/(1 + \mathrm{SNR})$. Unnormalized signal values depend on the RMS pattern contrast, which, in the present experiments, was fixed at 0.53. Therefore, $S^2 + N^2 = 0.53^2$ and the unnormalized signal values and SNR are related by $S^2 = 0.53^2\,\mathrm{SNR}/(1 + \mathrm{SNR})$.

5. One might think that direction repulsion affected subjects' judgments of perceived direction of motion when the stimuli were bidirectional. Direction repulsion effects in opponent (Marshak & Sekuler, 1979) and perpendicular (Hiris & Blake, 1996; Marshak & Sekuler, 1979) motion have been previously measured in dynamic random-dot stimuli. The effects, which amount to 4° or less, are too small, in our opinion, to influence the outcome of Experiment 3.

(Manuscript received September 23, 1996;
revision accepted for publication March 3, 1997.)