

Coherence and transparency of moving plaids composed of Fourier and non-Fourier gratings

JONATHAN D. VICTOR and MARY M. CONTE
Cornell University Medical College, New York, New York
and The Rockefeller University, New York, New York

We examined the perceptual coherence of two-component moving plaids. The gratings that constituted the plaids were either standard Fourier gratings (F), in which luminance was determined by a drifting sinusoid, or non-Fourier gratings (NF), in which the contrast of a random background was modulated by a drifting sinusoid. These NF gratings are examples of stimuli that generate a compelling percept of motion, even though they fail to elicit a motion signal from motion analyzers based on standard cross-correlation (Chubb & Sperling, 1988). Naive observers viewed three types of stimuli consisting of superpositions of these two components: (1) two standard drifting gratings (F/F), (2) two non-Fourier drifting gratings (NF/NF), and (3) one standard and one non-Fourier drifting grating (F/NF). As expected, the F/F stimulus yielded a compelling percept of coherent motion. The dominant percept of all the observers for the NF/NF stimulus was one of coherent motion, provided that both gratings were visible and of approximately equal contrast. None of the observers reported a dominant percept of coherent motion for the F/NF condition, over a wide range of contrasts for the two grating components and across two varieties of NF gratings. In view of the results of Albright (1992) and Albright and Chaudhuri (1989), that show that single cells in macaque V1 and MT respond to both F and NF motion, one cannot interpret our findings as evidence that F and NF motion are processed independently. Alternative, "higher level" interpretations based on the intrinsically ambiguous nature of the stimuli and physical laws governing the appearance of transparent objects are discussed.

The percept of motion is not an intrinsic aspect of a visual scene. Rather, this percept must be extracted by neural computations. For some simple stimuli, this computation selects a particular motion percept out of a range of possible percepts, each of which is physically consistent with the stimulus. For example, a drifting sinusoidal pattern is usually perceived as moving in a direction perpendicular to the bars of the grating. However, the physical stimulus is also consistent with simultaneous motion along (parallel to) the bars. Because the sine grating is homogeneous in this direction, such motion is invisible. A similar ambiguity applies to the interpretation of motion of an edge. This is known as the *aperture problem* (Nakayama, 1985).

Such ambiguities in the interpretation of motion are not generally manifest in the viewing of natural scenes, which are typically more complex than either sine gratings or edges. Thus, one suspects that the visual system combines the multiple motion cues present in a natural scene to disambiguate motion percepts. The perception of moving

plaid patterns provides a way to study the combination of such cues (Movshon, Adelson, Gizzi, & Newsome, 1985). A plaid pattern consisting of a superposition of two moving sine gratings is often interpreted as a moving grid. To a first approximation, the perceived direction of pattern motion is given by the unique direction that is compatible with each of its components—the intersection of constraints model of Adelson and Movshon (1982). More recently, studies with a wider range of plaid patterns have suggested deviations from this simple kind of interaction (Ferrera & Wilson, 1990).

A superposition of two gratings has a second interpretation, namely, that of two surfaces (the gratings) moving independently. A variety of stimulus parameters influence the observer's preference for the coherent (moving pattern) percept or the noncoherent percept. Gratings tend to be perceived as a coherent moving pattern if they are similar in spatial frequency and contrast (Adelson & Movshon, 1982), and tend to be perceived as separate surfaces if stimulus luminance values are consistent with physical transparency (Stoner, Albright, & Ramachandran, 1990).

In this context, we have chosen to study the interaction of motion signals generated by two kinds of grating patterns. The first kind of grating is a standard luminance grating. For standard luminance gratings, a variety of closely related computational strategies (Nakayama, 1985; Simoncelli & Adelson, 1991; van Santen & Sperling, 1985)

A portion of this work was presented at the 1991 meeting of the Association for Research in Vision and Ophthalmology, Sarasota, Florida. This work was supported in part by Grant EY7977 from the National Eye Institute (NIH). Address correspondence to J. D. Victor, Department of Neurology and Neuroscience, Cornell University Medical College, 1300 York Avenue, New York, NY 10021 (e-mail: victor@rockvax.bitnet).

suffice to extract motion signals. These include cross-correlation (Reichardt, 1961), detection of spatiotemporal gradients (Heeger, 1987), and spatiotemporal motion energy (Adelson & Bergen, 1985). Indeed, the properties of directionally selective single cells of primary visual cortex are consistent with a calculation of motion energy (Emerson, Bergen, & Adelson, 1992; Emerson, Citron, Vaughn, & Klein, 1987). These gratings will be called *Fourier* (F) motion stimuli, because their motion is manifest in their Fourier decomposition into sinusoidal luminance gratings.

The second kind of grating pattern is based on the *drift-balanced* motion stimuli of Chubb and Sperling (1988). These gratings generate strong percepts of motion, despite the failure of the aforementioned computational mechanisms to extract a motion signal. Because the motion percept generated by these patterns is not apparent from its Fourier decomposition into luminance gratings, we will call such stimuli *non-Fourier* (NF) motion stimuli. Chubb and Sperling (1989) suggest that a motion signal may be extracted from an NF motion stimulus by a computational mechanism that contains a local nonlinearity prior to standard motion analysis.

These considerations motivated us to examine the perceptual coherence of plaids composed of two NF gratings, and of plaids composed of one F grating and one NF grating. If NF motion were processed by a mechanism similar to the one that processes standard motion, then plaids composed of NF gratings should exhibit coherence analogous to that of standard plaids. Furthermore, if F and NF motion signals are carried by the same neurons, then a "low-level" account of grating coherence would predict coherence of mixed F/NF gratings. Although we find robust coherence of NF gratings, we find no evidence for coherence of mixed gratings. This is difficult to account for on the basis of low-level mechanisms, and we therefore consider high-level explanations.

METHODS

Visual Stimuli

We begin with a formal description of the visual stimuli and some of their properties. Spatial position will be denoted by the vector quantity \mathbf{x} . The two coordinates of \mathbf{x} are restricted to be integers, so that \mathbf{x} specifies a pixel. The luminance of the stimulus at position \mathbf{x} and time t will be denoted $L(\mathbf{x}, t)$. For all stimuli, luminance is given by a mean (background) luminance L_0 modulated by two spatiotemporal patterns $P_i (i=1,2)$, each at contrast c_i . That is,

$$L(\mathbf{x}, t) = L_0 [1 + c_1 P_1(\mathbf{x}, t) + c_2 P_2(\mathbf{x}, t)]. \quad (1)$$

The modulation pattern P for a standard (Fourier) drifting sinusoidal grating of spatial frequency $\mathbf{k}=(k_1, k_2)$ and temporal frequency f is given by

$$P_{F, \mathbf{k}, f}(\mathbf{x}, t) = \cos[2\pi(\mathbf{k} \cdot \mathbf{x} - ft)]. \quad (2)$$

The speed of this grating $|\mathbf{v}|$ is $f/|\mathbf{k}|$; its (vector) velocity \mathbf{v} is $\mathbf{v} = \mathbf{k}f/|\mathbf{k}|^2$.

Our main non-Fourier motion stimulus was based on a modulation pattern P , given by

$$P_{NF, \mathbf{k}, f}(\mathbf{x}, t) = \cos\{\pi[R(\mathbf{x}) + \mathbf{k} \cdot \mathbf{x} - ft]\}. \quad (3)$$

Here, $R(\mathbf{x})$ is a random function that assigns the value 0 or 1 to each pixel \mathbf{x} independently and with equal probability. Note that only a factor of π , and not 2π , appears in the argument of the cosine function in Equation 3. Since the phase shift induced by $R(\mathbf{x})$ amounts to a random inversion of polarity at each pixel, pairwise correlations between individual pixels are zero (averaged over the ensemble of choices for the random function R). Although this is distinct from the examples of non-Fourier motion stimuli presented by Chubb and Sperling (1988), it shares the critical property of being microbalanced (and therefore its motion is not detectable by Reichardt-type cross-correlators or Fourier energy models).

As suggested by Chubb and Sperling (1988), the motion in the pattern $P_{NF, \mathbf{k}, f}(\mathbf{x}, t)$ becomes manifest if it is subjected to a local nonlinearity (for example, a square law):

$$[P_{NF, \mathbf{k}, f}(\mathbf{x}, t)]^2 = \frac{1}{2} + \frac{1}{2} \cos[2\pi(\mathbf{k} \cdot \mathbf{x} - ft)]. \quad (4)$$

Thus, the squared value is a standard grating superimposed on an offset of 1/2.

At any fixed position, the modulation pattern $P_{NF, \mathbf{k}, f}$ varies sinusoidally in time, at a frequency $f/2$ (Equation 3), while its square (Equation 4) is modulated sinusoidally with frequency f .

These stimuli were weighted with appropriate contrasts and combined by superposition according to Equation 1. Superposition of two Fourier gratings (Equation 2) forms a standard plaid, of the sort studied extensively by Movshon et al. (1985), Ferrera and Wilson (1990), and others. We will denote this stimulus by F/F. Superposition of two non-Fourier gratings (Equation 3) forms a stimulus we will denote by NF/NF; superposition of one Fourier grating (Equation 2) and one non-Fourier grating (Equation 3) forms a heterogeneous plaid, which we will denote by F/NF. Figure 1 shows frames of stimuli constructed from a single Fourier grating (Equation 2), a single non-Fourier grating (Equation 3), and the three types of plaids. These stimuli were used in Experiment 1.

In addition, we generated another kind of non-Fourier motion pattern, similar to that used by Stoner and Albright (1992) and Albright and Stoner (1989). The initial frame of the pattern $P_{NF', \mathbf{k}, f}(\mathbf{x}, 0)$ consists of a random assignment of each pixel to states +1 or -1. On successive frames (four raster refreshes at 135 Hz, or 29.6 msec), some pixels changed states according to a random function $U_{\mathbf{k}, f}(\mathbf{x}, t)$, where $U_{\mathbf{k}, f}(\mathbf{x}, t) = 1$ indicates no change in state, and $U_{\mathbf{k}, f}(\mathbf{x}, t) = -1$ indicates a flip in state. That is,

$$P_{NF', \mathbf{k}, f}(\mathbf{x}, t+1) = U_{\mathbf{k}, f}(\mathbf{x}, t) \cdot P_{NF', \mathbf{k}, f}(\mathbf{x}, t). \quad (5)$$

The motion is carried by the random function $U_{\mathbf{k}, f}(\mathbf{x}, t)$:

$$\text{prob}\{U_{\mathbf{k}, f}(\mathbf{x}, t)=1\} = p_0 + p_m \cos[2\pi(\mathbf{k} \cdot \mathbf{x} - ft)]. \quad (6)$$

That is, the probability that a pixel changes state from one frame to the next is a sinusoidal function of position, of mean p_0 , and modulation depth p_m . In these studies, we chose $p_0 = p_m = .125, .25$, or $.5$. The ratio p_m/p_0 corresponds to the "flicker contrast" parameter of Stoner and Albright (1992).

The appearance of this stimulus was that of a random-dot field with traveling waves of "snow," which is closely related to Chubb and Sperling's (1988) Example D of their Figure 4. Extraction of motion from this stimulus can be accomplished by temporal filtering, followed by a local nonlinearity, followed by cross-correlation. Because any individual frame consisted of pixels assigned to the states +1 and -1 in a random and uncorrelated fashion, the motion cannot be extracted merely by a local nonlinearity followed by cross-correlation; the initial temporal filtering is required (Chubb & Sperling, 1988). We created heterogeneous plaids, denoted F/NF', by superimposing the Fourier grating pattern (Equation 5) on the non-Fourier grating pattern (Equation 2). These stimuli were used in Experiment 2. Here again, the formal definition of superposition is Equation 1.

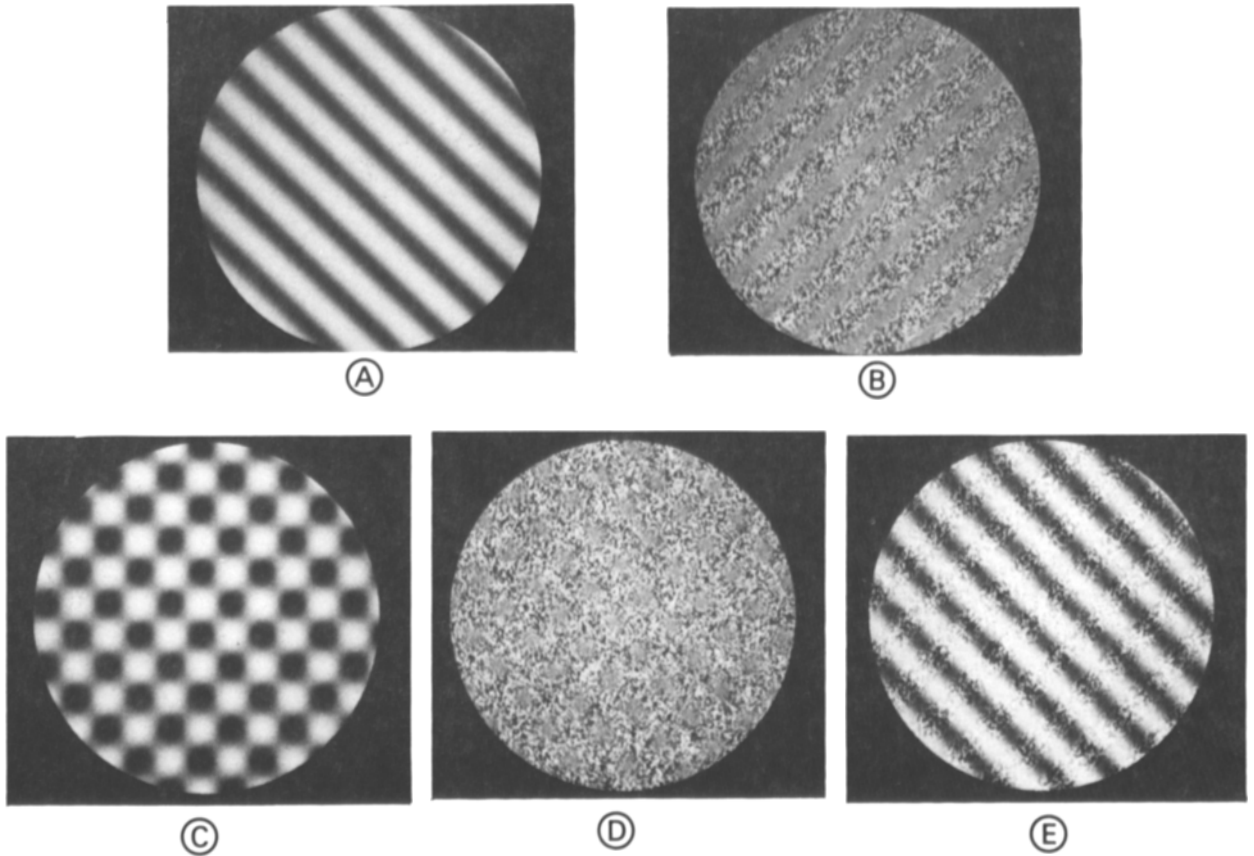


Figure 1. Photographs of visual stimuli. (A) A standard (Fourier) grating. (B) A non-Fourier grating. (C) A standard (F/F) plaid. (D) A non-Fourier (NF/NF) plaid. (E) A heterogeneous (F/NF) plaid.

The stimuli described above were produced on a Conrac 7351 monitor. R, G, and B intensity signals for the 135 Hz raster display were generated by specialized electronics (Milkman et al., 1980) interfaced to a DEC 11/73 computer. The electronics included a 12-bit look-up table to linearize the intensity/voltage characteristics of the CRT. Each pixel of the 256×256 pixel display subtended 3.2 min. The display was viewed binocularly at a distance of 114 cm and had a mean luminance of 50 cd/m^2 ; a circular mask, 14° diam, was placed immediately in front of the display. The grating spatial frequency was 0.57 c/deg and drift velocity was 0.92 deg/sec .

Subjects

The subject pool consisted of 5 volunteers, 2 male and 3 female. All but 1 subject (M.C.) were naive to the purpose of the experiments. The nonnaive subject (M.C.) participated only in Experiment 1. The subjects ranged in age from 26 to 41 years, and were corrected to normal visual acuity, if necessary.

Psychophysical Methods

The plaid stimuli were constructed from pairs of gratings whose drift directions differed by 90° and were presented for periods of 5 sec. The subjects were asked to "indicate the dominant direction or directions of motion" to within 45° through the use of the terms *cardinal* or *diagonal*. Reports of a single dominant drift direction along the direction of pattern motion were recorded as a percept of coherence. Reports of either a single dominant drift direction in the direction of one of the component gratings or two directions

of motion were recorded as a percept of noncoherence. No other reports (e.g., three directions of motion) were obtained. We refrained from using terms such as *coherence*, *plaid*, or *transparency* with the naive subjects. An equal number of stimuli were presented with each of eight directions (four cardinal, four oblique) of drift; the order of presentations was randomized. Presentations of different contrasts were intermixed randomly.

In a control experiment, the subjects were asked to indicate (forced choice) whether one or two gratings were perceived. One half of the trials consisted of a moving high-contrast grating; the other half consisted of a plaid, in which the second grating was present at low contrast. This control experiment followed the coherence judgments.

RESULTS

Experiment 1

We determined the frequency of the coherent percept for a binary superposition as a function of the contrasts of its two components. One grating component was presented at contrasts (c_1 in Equation 1) that varied from 0.05 to 0.5; the second component was presented at contrasts (c_2 in Equation 1) of either 0.17 or 0.35. Figure 2A shows the results obtained for a superposition of two Fourier gratings P_F . Coherence was the dominant percept in all

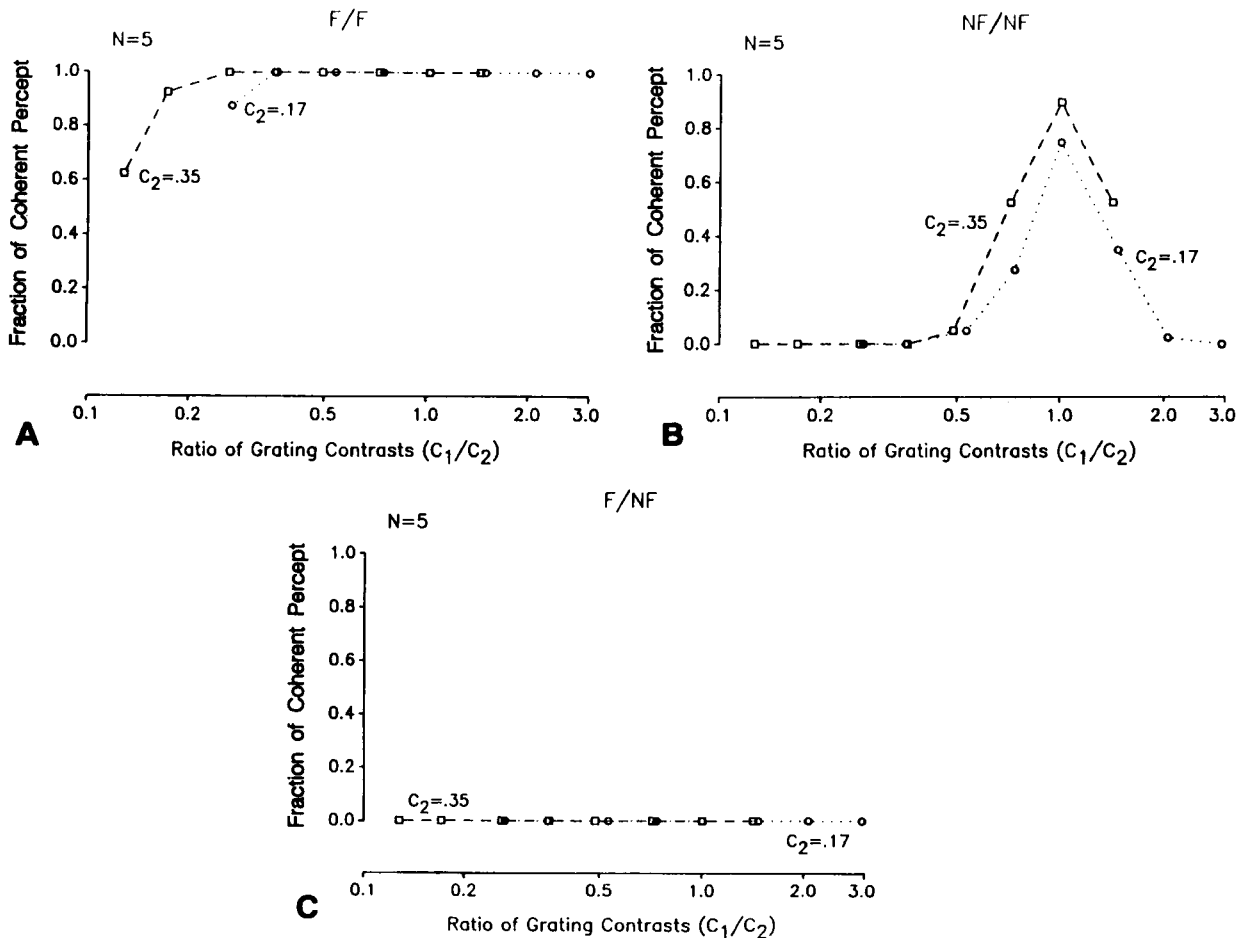


Figure 2. Psychophysical responses to moving plaids. (A) Standard (F/F) plaid. (B) Non-Fourier (NF/NF) plaid. (C) Mixed (F/NF) plaid. In the mixed-plaid condition, c_1 denotes the contrast of the Fourier component and c_2 denotes the contrast of the non-Fourier component. Fractions of coherent percepts are averaged across 5 subjects.

cases, and was the exclusive percept, provided that the two contrasts did not differ by more than a factor of 3.

Figure 2B shows the results obtained for two non-Fourier gratings of type P_{NF} (Equation 3). Coherence was the dominant percept when the components were of equal contrast ($c_1/c_2=1$), and coherence was not perceived for contrasts that differed by more than a factor of 2. For a contrast ratio of 1.4 ($c_1/c_2=1.4$ or $c_2/c_1=1.4$), coherence was perceived approximately half of the time. For any given contrast ratio, coherence was perceived more frequently at higher contrast.

In the equal-contrast conditions, one subject (N.S.) perceived coherence in 16/16 presentations. One subject (W.C.) perceived coherence in 7/16 presentations; other subjects' responses were between these extremes. In the conditions with a contrast ratio of 1.4, frequency of the coherent percept ranged from 25/32 (N.S.) to 8/32 (M.I. and S.L.). In the conditions with a contrast ratio of 2.0, frequency of the coherent percept ranged from 10/32 (N.S.) to 0/32 (S.L.). Thus, although there was a signif-

icant variation across subjects in the overall frequency of the coherent percept, all the subjects readily reported coherence in the equal-contrast condition and progressively shifted their responses away from coherence for contrast ratios other than 1.

We noted no bias of the perceived direction of motion toward cardinal or oblique directions. Furthermore, if such a bias were present, the balanced presentation scheme (see Methods) ensured that it would not influence the frequency of the coherent percept.

Figure 2C shows performance for a superposition of one Fourier grating P_F and one non-Fourier grating P_{NF} . In this mixed condition, none of the subjects perceived coherence in any trial at any contrast ratio. Most of the subjects volunteered that they perceived two directions of motion over most of the contrast range. Following the experiment, we asked the subjects to detect the presence of a low-contrast Fourier grating in the presence of a high-contrast ($c_2=0.17$ or $c_2=0.35$) masking grating of type NF. All the subjects showed perfect performance in eight

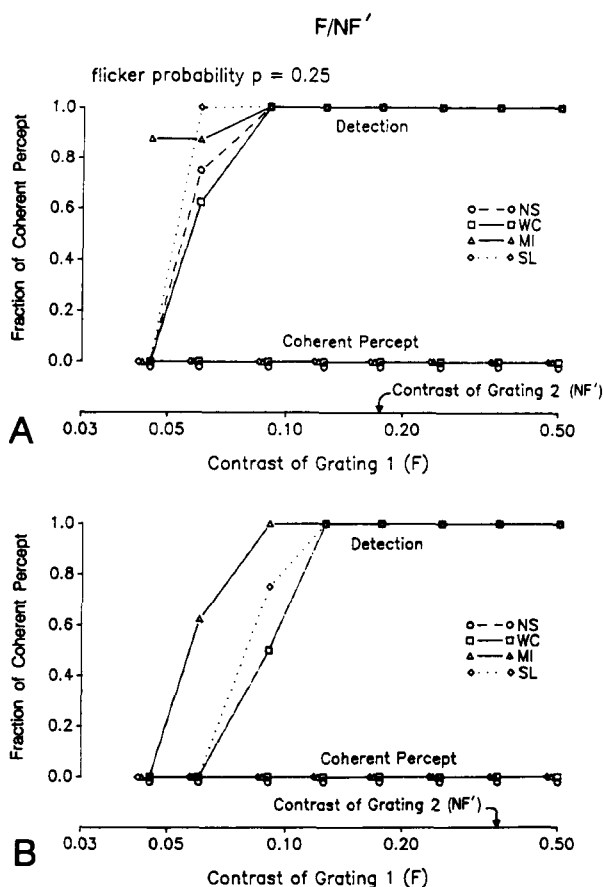


Figure 3. Psychophysical responses to mixed plaids of type F/NF' . Flicker probability mean p_0 and modulation depth p_m are both equal to .25. (A) NF' component at contrast $c_2 = 0.17$. (B) NF' component at contrast $c_2 = 0.35$.

forced-choice presentations, in which c_1 was either 0.03 or 0. Thus, failure to perceive coherence was not related to failure to perceive the low-contrast grating.

Experiment 2

In this experiment, we measured the frequency of coherence for the mixed F/NF' plaids (Albright & Stoner, 1989; Stoner & Albright, 1992). Figure 3 shows the results for NF' gratings constructed with a mean flicker and flicker probability modulation depth p_m of .25 (Equation 6). The contrast of the NF' grating was 0.17 (see Figure 3A) or 0.35 (see Figure 3B). No subject perceived coherence during any of the presentations of F/NF' plaids, for any contrast of the Fourier or non-Fourier components. To be sure that both gratings were visible, the subjects also performed a detection task, as described in Methods. At low contrasts, most of the subjects were not able to detect the Fourier grating. However, masking of the Fourier grating by the NF' grating cannot explain the lack of a coherent percept at the higher contrasts of the Fourier grating.

Figure 4 shows the results pooled across subjects for flicker probabilities $p_0 = p_m = .125, .25, \text{ and } .5$. Again, there was never a report of a coherent percept. The masking efficacy of the non-Fourier grating was similar for all flicker probabilities, and does not explain the lack of a coherent percept.

GENERAL DISCUSSION

Summary of Results

We found that plaids constructed from moving NF gratings elicit a percept of coherent motion, analogous to the coherent percept elicited by standard plaids (Adelson & Movshon, 1982). However, for NF/NF plaids, coherence was seen over a narrower range of relative contrasts than it was for F/F plaids (see Figures 2A and 2B). We also found the striking result that two kinds of mixed plaids (F/NF and F/NF') never elicited a coherent percept (see Figures 2C, 3, and 4).

This finding disagrees with that previously reported by Albright and Stoner, 1989 and Stoner and Albright, 1992 for the F/NF' plaids. The luminance contrast, flicker contrast, spatial parameters, and temporal parameters of the stimulus we used were similar to those used by these authors. Although coherence may well critically depend on relative grating contrast, this is an unlikely explanation for our failure to observe coherence. Stoner and Albright (1992) reported that the dependence of the fraction of reports of coherence on relative grating contrast had a full width at a half-maximum of 1.37–2.19 octaves in 3 of 4 subjects. Our sampling of relative grating contrasts by half-octaves (.15 log units) would not have missed the coherent percept if it were present.

In the mixed F/NF and F/NF' plaids, the NF grating was at high contrast and the F grating varied from $2\times$ to at least $10\times$ threshold, so our results cannot be explained by a failure to detect the F grating. More likely, the difference in results is due to the susceptibility of coherence judgments to subjective factors. Our results were obtained from subjects who were asked to indicate the dominant direction (or directions) of movement. These subjects were unpracticed in coherence judgments, and (except for M.C.) were not introduced to the notion of coherence in "practice" trials. In the studies of Albright and Stoner (1989) and Stoner and Albright (1992), subjects were first trained on a series of standard F/F plaids, and were explicitly instructed to judge "coherence." Although their studies indicate that a coherent interpretation of F/NF' gratings is possible, we conclude that the coherence of heterogeneous-cue plaids (F/NF or F/NF') is far weaker than the coherence of F/F or NF/NF plaids.

Comparison of Fourier and Non-Fourier Motion Extraction

Chubb and Sperling (1989) initially proposed that a standard computational mechanism that extracts F motion signals would also extract NF motion signals if the local luminance or signed contrast signal were transformed by

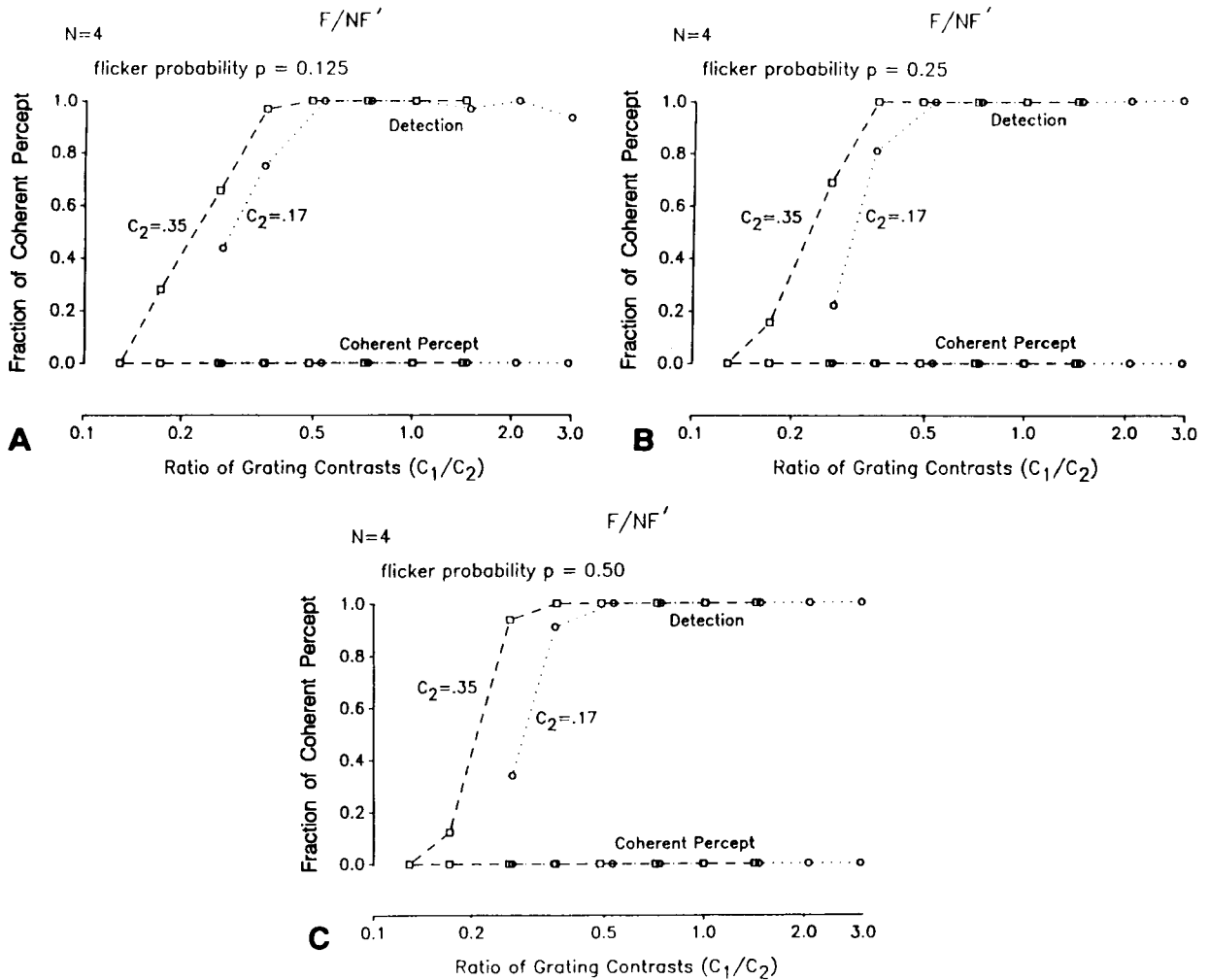


Figure 4. Psychophysical responses to mixed plaids of type F/NF'. Flicker probability mean p_0 and modulation depth p_m are both equal to .125 (A), .25 (B), and .5 (C). Fractions of coherent percepts are averaged across the 4 naive subjects. The data for B were taken from Figure 3.

a nonlinear function, provided that the nonlinear function contained even-order components. If this local nonlinearity contained odd-order components as well, then a single circuit could account for extraction of both kinds of motion.

Several lines of evidence suggest that F and NF motion are extracted by the same pathway. Velocity discrimination judgments for moving F and NF gratings are very similar (Turano & Pantle, 1989), and adaptation experiments provide additional psychophysical evidence for a common neural substrate (Turano, 1991). We (Victor & Conte, 1992) recently showed that the VEPs elicited by the onset of F and NF motion are nearly identical in contrast dependence, dynamics, velocity tuning, and scalp topography. Albright and Chaudhuri (1989) have shown that striate neurons that respond to F motion in a directionally selective manner typically respond to NF motion in a similar fashion.

Local nonlinearities typically generate "distortion products"—spatial frequencies consisting of integer linear combinations (e.g., $n_1 f_1 + n_2 f_2$) of spatial frequencies present in the untransformed stimulus. Distortion products alone do not account for the synthesis of the component velocity vectors of a plaid stimulus into the pattern velocity vector. This is because velocities and spatial frequencies are reciprocally related and, consequently, the pattern velocity vector generally does not correspond to any particular distortion product.

Possible Models

If F and NF motion signals are extracted by the same pathways, how can one account for the minimal coherence of F/NF plaids? It might be argued that the F/NF stimulus was never seen as coherent because the perceived contrasts of the gratings were never sufficiently close. Comparison of contrasts of the gratings in terms of mul-

tuples of their thresholds is not the appropriate way to settle this issue, because of contrast constancy at suprathreshold levels (Georgeson & Sullivan, 1975). Indeed, at the lowest F/NF contrast ratio (F contrast 0.05, NF contrast 0.35), the NF grating was sufficiently more salient than the F grating that the dominant effect was masking of the F grating by the NF grating, even though both gratings were approximately $10\times$ their thresholds when presented alone. At the highest F/NF contrast ratio (F contrast 0.50, NF contrast 0.17), the F grating (approximately $100\times$ threshold in isolation) was strikingly more salient than the NF grating (approximately $5\times$ threshold in isolation). Thus, although one cannot be sure which physical contrast ratio resulted in perceptually equal contrasts of the F and NF gratings, it is clear that this range was bracketed and (see above) adequately sampled.

An ingenious account of the minimal coherence of F/NF plaids was proposed by Wilson (1991). Wilson proposed motion analysis via banks of parallel computational units. Each computational unit in the model is tuned to a band of spatial frequencies. Effectively, the computational unit extracts both F and NF motion signals from a band-limited image of the stimulus. Interactions within a unit generate the pattern-motion direction of plaids. In this model, computational units tuned to different spatial frequencies do not interact. This accounts for the familiar finding (Adelson & Movshon, 1982) that F gratings fail to cohere if their spatial frequencies are sufficiently different.

To analyze the behavior of the model for NF gratings, it is helpful to consider the frequency-domain consequences of multiplication of a low-frequency signal (the "envelope") by a second broadband or high-frequency signal (the "carrier"). This is the essence of the non-Fourier motion stimulus (Equation 3): a drifting sinusoidal envelope multiplied by a random function $R(\mathbf{x})$ of space. The frequency content of the product consists of all sums and differences of the frequencies present in the envelope and the carrier. If the frequencies in the envelope are much lower than the frequencies present in the carrier, then the frequencies in the resulting product remain similar to those in the carrier, even though there is a percept of a grating at the frequency of the envelope.

Because the Wilson model postulates interactions on the basis of the stimulus frequency alone, it accounts for the failure of F and NF gratings to cohere. This is because the NF grating is essentially an amplitude-modulated grating with a broadband carrier, and has very little power at the spatial frequency of the F grating. The Wilson model also accounts for the observed coherence of two NF gratings, because their Fourier spectra are spread over the same range. Interestingly, the Wilson model predicts that F and NF gratings will cohere if the carrier of the NF grating is matched to the spatial frequency of the F grating. We have set up several stimuli of this type, and have observed no hints of coherence. However, the parameter space of such stimuli is very large, and this negative finding does not necessarily defeat the model.

In general, a low-level explanation of our results seems unlikely, in that it is difficult to envision how F and NF signals could be separated once they are generated by the same circuits. On the other hand, higher level processes, such as shading consistent with physical transparency, influence the perceptual coherence of an otherwise ambiguous plaid (Stoner & Albright, 1991; Stoner et al., 1990). For even a single grating, it is well known that the perceived direction of movement may be influenced by the shape of the aperture through which it is viewed (Movshon et al., 1985). We therefore turn to possible higher-level accounts of our results.

Higher Level Processes

We begin with the notion that a standard F/F plaid stimulus may be interpreted either as a pair of surfaces, each moving in its "component" direction, or as a single surface, moving in the "pattern" direction. The pair-of-surfaces interpretation is consistent with the physics of transparency only if the luminance values obey certain algebraic relations (Kersten, 1990). A plaid composed of two-value gratings of contrasts c_a and c_b has four possible luminance values, one for each possible choice of values from each component (see Figure 5A). To be consistent with pure transparency (Stoner et al., 1990), the four luminance values of the plaid must obey

$$I_{01}I_{10} - I_{00}I_{11} = 0. \quad (7)$$

A plaid stimulus constructed of additively superimposed gratings of luminance L_0 and contrasts c_a and c_b (see Figure 5B) has luminance values given by

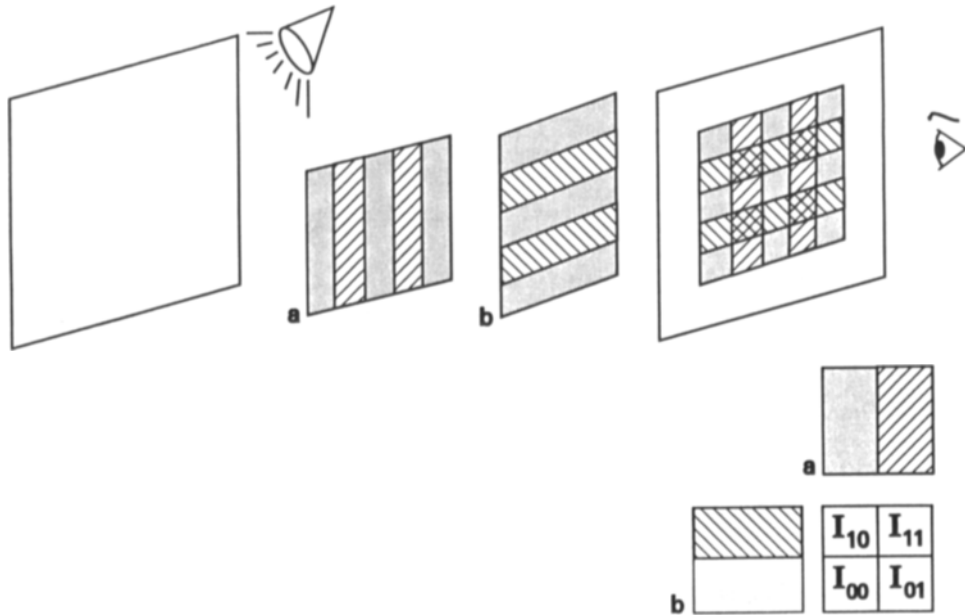
$$\begin{aligned} I_{11} &= L_0(2 + c_a + c_b), \\ I_{10} &= L_0(2 + c_a - c_b), \\ I_{01} &= L_0(2 - c_a + c_b), \\ I_{00} &= L_0(2 - c_a - c_b). \end{aligned} \quad (8)$$

In general, the luminance values generated by superposition (Equation 8) are inconsistent with the requirements of pure transparency (Equation 7). The discrepancy is $4L_0^2c_ac_b$. That is, at any given overall contrast level $c_a + c_b$, the plaid is most inconsistent with transparency when c_a and c_b are comparable, and least inconsistent with transparency when the contrast ratio is large. This provides an alternative interpretation for the observation (Adelson & Movshon, 1982) that coherence is strongest when grating contrasts are comparable—this is the condition which is least consistent with the physics of transparency.

A similar line of reasoning can be applied to plaids composed of NF components. For simplicity, we consider a plaid whose bars change abruptly from high contrast to low contrast, rather than the smoothly varying stimuli used in the above experiments. The NF/NF plaid has a potential interpretation as either a pair of transparent surfaces (see Figure 6A) or a single surface (see Figure 6B). A

(A)

TRANSPARENCY, F/F



(B)

SUPERPOSITION, F/F

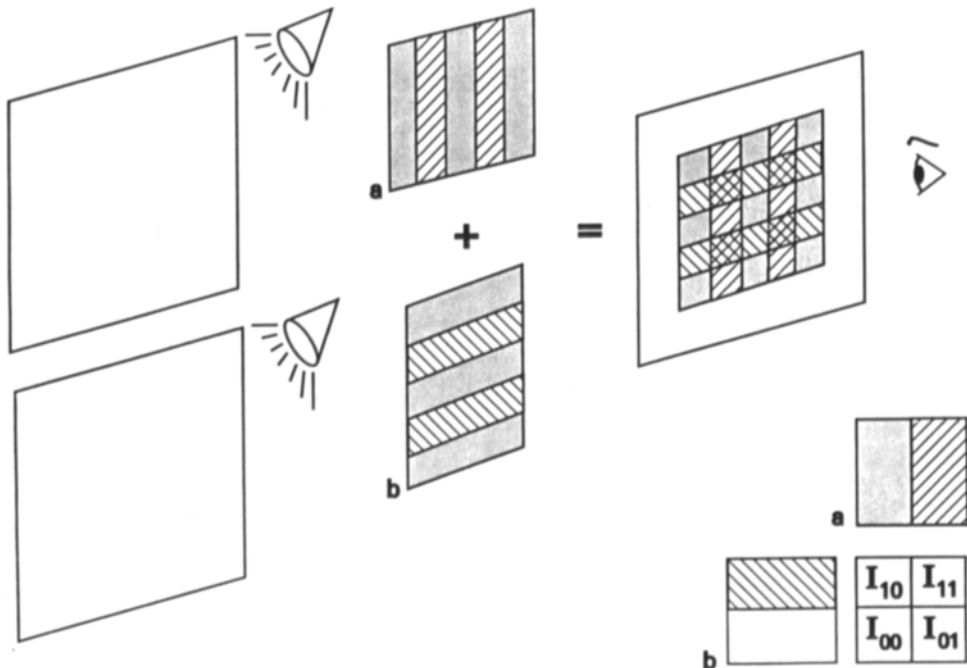


Figure 5. (A) The transparent interpretation of an F/F plaid stimulus. (B) Superposition of two F gratings to form a plaid.

region of high local contrast represents the appearance of a speckled background through a clear (nondiffusing) section of a grating; a uniform (zero contrast) region represents the appearance of this background through a blurring (diffusing) section of the grating. The transparency construct (see Figure 6A) would imply that three quarters of the plaid is at zero contrast—the three quarters that are seen through the diffusing sections of *either* component grating. Superposition, however, creates a plaid only one quarter of which is at zero contrast—that quarter that receives a zero-contrast contribution from *both* gratings (see Figure 6B). Again, the superposition of stimuli is inconsistent with the physics of transparency, and again, the coherent interpretation is seen.

The mixed (F/NF) plaid is analyzed in Figure 7. Surface a, which corresponds to the F grating, subdivides the image into vertical strips of alternate intensities I_0 and I_1 (see Figure 7A). Surface b, which corresponds to the NF grating, alternately subdivides each of these strips into regions of high local contrast (clear portions of b) and zero local contrast (blurring portions of b). Thus, the physics of transparency implies that there will be four kinds of regions in the image (see Figure 7A): two of intensity I_0 , two of intensity I_1 , and two with high local contrast, two with zero local contrast.

The superposition of crossed F and NF gratings (Figure 7B) generates a plaid that contains four kinds of regions: two of high intensity and two of low intensity, and two with high local contrast and two with zero local contrast. These four regions have the same characteristics and configuration as the regions generated by the corresponding transparent surfaces (Figure 7A) do. Thus, only in the F/NF case is the stimulus (a superposition) consistent with the physical constraints of transparency. For F/F and NF/NF plaids, the stimulus is inconsistent with the physical properties of real transparent objects. We therefore proposed (Victor & Conte, 1991) that this is an important factor in the perception of noncoherence in the F/NF stimulus and coherence in the F/F and NF/NF stimuli.

There are problems with this model. A minor problem is that in the pure-transparency interpretation of the F/NF stimulus, the local Weber contrast of the nonblurred region is the same in both the high-intensity portion (I_0) and the low-intensity portion (I_1). But in the superposition stimulus, there is a slightly lower local Weber contrast in the high-intensity region. This is because of greater “dilution” of local contrast by the more intense light originating from the F grating. However, it is unclear whether the visual system is able to detect these small differences in local contrast, or whether local-contrast cues are utilized in assessing surface properties such as transparency.

There are more serious problems with this model. The algebraic constraint (Equation 7) holds only for pure

transparency. Luminance values consistent with surfaces that are partially transparent (and partially reflective) may be generated by superposition (Equation 8), and such F/F plaids may nevertheless be seen as coherent.

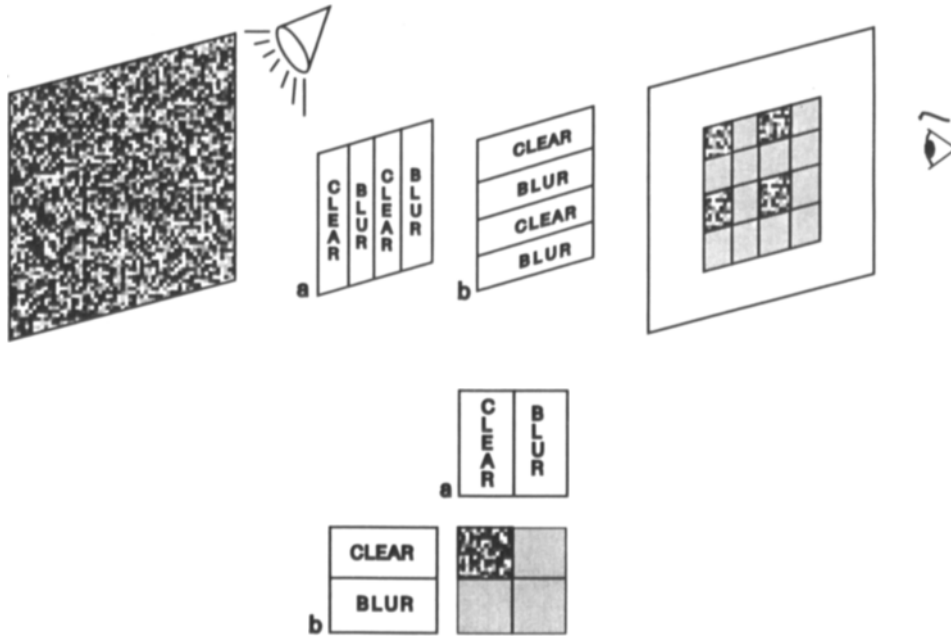
Conversely, this model predicts that plaids such as those of Figure 6A (which can be generated by multiplicative combination of two non-Fourier gratings) will be seen as noncoherent, because they are consistent with transparency. This is not the case; instead, this stimulus generates the percept of moving patches of contrast on a uniform background, rather than that of two sliding blurring surfaces. This implies that other factors are likely at work.

One such factor may be a Bayesian process, in which an intrinsically ambiguous percept is most likely to be interpreted as an object in generic view, rather than as an object that must be oriented or illuminated in a particular fashion (Nakayama & Shimojo, 1991). For example, in the F/NF case (Figure 7), both coherence and transparency were consistent with the stimulus. Coherence requires that the surface have four kinds of regions, with local contrast and luminance varying in a prescribed, non-generic fashion. Transparency requires no particular relationship between the properties of the perceived surfaces. Thus, the Bayesian principle leads to a percept of transparency. If a plaid stimulus is inconsistent with the transparent interpretation, as in the F/F case (Figure 5) or the NF/NF case (Figure 6), then no Bayesian inference need be made; in these cases, coherent motion is the only percept. Finally, a plaid stimulus constructed to match that of the transparent interpretation of the NF/NF stimulus (Figure 6A) would still be viewed as coherent, since the transparent interpretation requires a pair of *matched* orthogonal surfaces.

A general premise of any higher level interpretation of our findings is that initial motion processing provides cues that may be interpreted in more than one way—in the present case, coherent or noncoherent motion—and that a selection process is required at a later stage. Whatever its mechanism, there is evidence that this selection process occurs in MT and its human analog. In the plaid paradigm, some MT units respond both to the component-motion direction and to the plaid motion direction, in a manner that parallels perceptual coherence (Stoner & Albright, 1991). MT also appears to be critical in detecting a pattern of global motion from a field of moving dots (Newsome, Britten, & Movshon, 1989; Newsome & Paré, 1988; Salzman, Britten, & Newsome, 1990). The use of this motion information for structure from motion (Siegel & Andersen, 1986, 1988) requires the implicit selection between the percept of random motion and the percept of motion in two opposing directions within separate planes. In the context of what extrastriate motion analysis does, it thus makes sense to consider coherence versus noncoherence for plaids to be a matter of selection, rather than merely calculation.

(A)

TRANSPARENCY, NF/NF



(B)

SUPERPOSITION, NF/NF

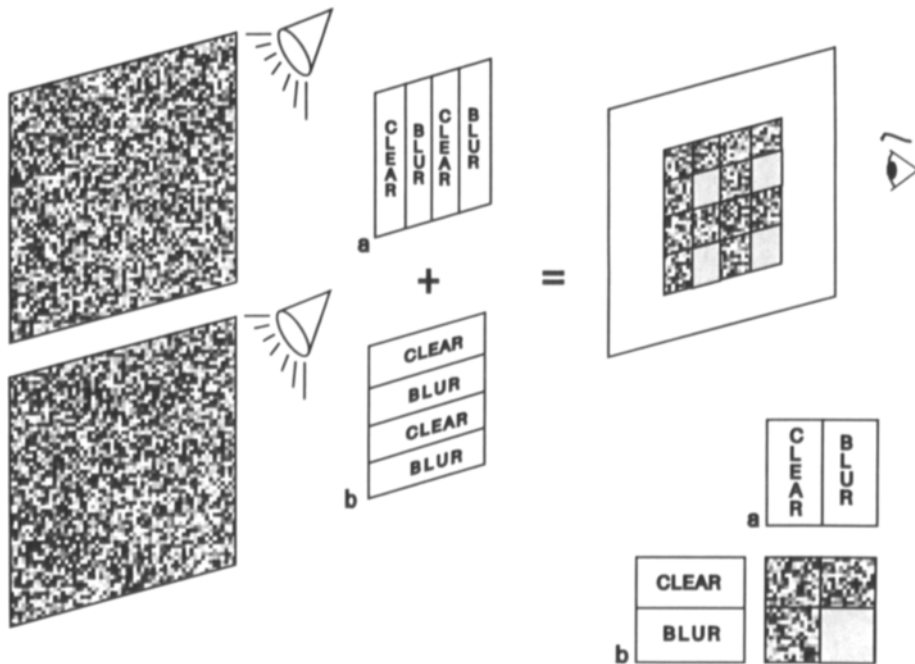
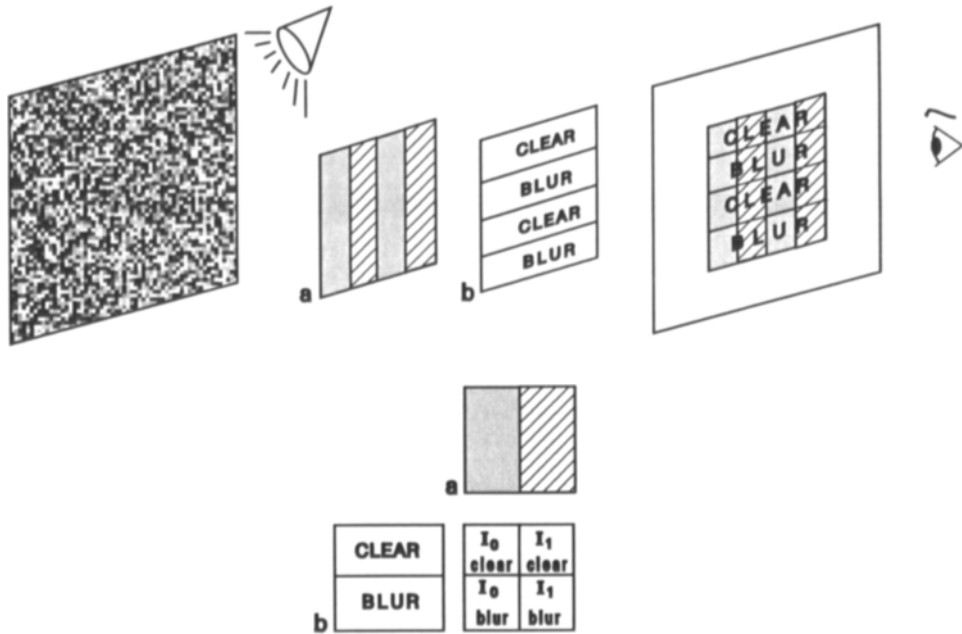


Figure 6. (A) The transparent interpretation of an NF/NF plaid stimulus. (B) Superposition of two NF gratings to form a plaid.

Ⓐ

TRANSPARENCY, F/NF



Ⓑ

SUPERPOSITION, F/NF

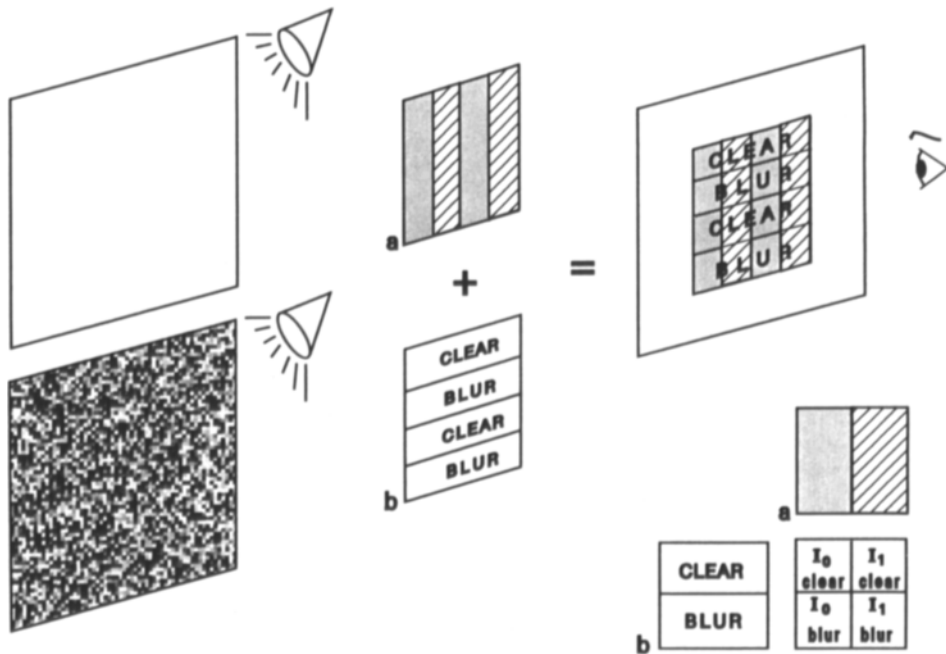


Figure 7. (A) The transparent interpretation of a mixed F/NF plaid stimulus. (B) Superposition of F and NF gratings to form a plaid.

REFERENCES

- ADELSON, E. H., & BERGEN, J. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America*, **A2**, 284-299.
- ADELSON, E. H., & MOVSHON, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, **300**, 523-525.
- ALBRIGHT, T. D. (1992). Form-cue invariant motion processing in primate visual cortex. *Science*, **255**, 1141-1143.
- ALBRIGHT, T. D., & CHAUDHURI, A. (1989). Orientation selective responses to motion contrast boundaries in macaque V1. *Neuroscience Abstracts*, **15**, 323.
- ALBRIGHT, T. D., & STONER, G. R. (1989). Motion perception survives figural cue heterogeneity. *Investigative Ophthalmology & Visual Science*, **30**(Suppl.), 74.
- CHUBB, C., & SPERLING, G. (1988). Drift-balanced random stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America*, **A5**, 1986-2006.
- CHUBB, C., & SPERLING, G. (1989). Two motion perception mechanisms revealed through distance-driven reversal of apparent motion. *Proceedings of the National Academy of Science*, **86**, 2985-2989.
- EMERSON, R. C., BERGEN, J. R., & ADELSON, E. H. (1992). Directionally selective complex cells and the computation of motion energy in cat visual cortex. *Vision Research*, **32**, 203-218.
- EMERSON, R. C., CITRON, M., VAUGHN, W. J., & KLEIN, S. (1987). Nonlinear directionally selective subunits in complex cells of cat striate cortex. *Journal of Neurophysiology*, **58**, 33-65.
- FERRERA, V. P., & WILSON, H. R. (1990). Perceived direction of moving two-dimensional patterns. *Vision Research*, **30**, 273-287.
- GEORGEON, M. A., & SULLIVAN, G. D. (1975). Contrast constancy: Deblurring in human vision by spatial frequency channels. *Journal of Physiology*, **252**, 627-656.
- HEEGER, D. (1987). Model for the extraction of image flow. *Journal of the Optical Society of America*, **A4**, 1455-1471.
- KERSTEN, D. (1990). Transparency and the cooperative computation of scene attributes. In M. S. Landy & J. A. Movshon (Eds.), *Computational models of visual processing* (pp. 209-228). Cambridge, MA: MIT Press.
- MILKMAN, N., SCHICK, G., ROSSETTO, M., RATLIFF, F., SHAPLEY, R., & VICTOR, J. D. (1980). A two-dimensional computer-controlled visual stimulator. *Behavior Research Methods & Instrumentation*, **12**, 283-292.
- MOVSHON, J. A., ADELSON, E. H., GIZZI, M. S., & NEWSOME, W. T. (1985). The analysis of moving visual patterns. In C. Chagas, R. Gattass, & C. Gross (Eds.), *Pattern recognition mechanisms* (pp. 117-151). *Experimental Brain Research* (Suppl. 11). Berlin: Springer-Verlag.
- NAKAYAMA, K. (1985). Biological image motion processing: A review. *Vision Research*, **25**, 625-660.
- NAKAYAMA, K., & SHIMOJO, S. (1991). Bayesian inference and the perception of untextured stereograms. *Investigative Ophthalmology & Visual Science*, **32**(Suppl.), 696.
- NEWSOME, W. T., BRITTEN, K. H., & MOVSHON, J. A. (1989). Neuronal correlates of a perceptual decision. *Nature*, **341**, 52-54.
- NEWSOME, W. T., & PARÉ, E. B. (1988). A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *Journal of Neuroscience*, **8**, 2201-2211.
- REICHARDT, W. (1961). Autocorrelation, a principle for the evaluation of sensory information by the central nervous system. In W. A. Rosenbluth (Ed.), *Sensory communication* (pp. 303-317). New York: Wiley.
- SALZMAN, C. D., BRITTEN, K. H., & NEWSOME, W. T. (1990). Cortical microstimulation influences perceptual judgments of motion direction. *Nature*, **346**, 174-177.
- SEGEL, R. M., & ANDERSEN, R. A. (1986). Motion perceptual deficits following ibotenic acid lesions of the middle temporal area (MT) in the behaving rhesus monkey. *Society for Neuroscience Abstracts*, **12**, 1183.
- SEGEL, R. M., & ANDERSEN, R. A. (1988). Perception of three-dimensional structure from motion in monkey and in man. *Nature*, **331**, 259-261.
- SIMONCELLI, E. P., & ADELSON, E. H. (1991). Relationship between gradient, spatiotemporal-energy, and regression models for motion perception. *Investigative Ophthalmology & Visual Science*, **32**(Suppl.), 893.
- STONER, G. R., & ALBRIGHT, T. D. (1991). Responses of area MT neurons to non-coherently moving plaid patterns. *Investigative Ophthalmology & Visual Science*, **32**(Suppl.), 822.
- STONER, G. R., & ALBRIGHT, T. D. (1992). Motion coherency rules are form-cue invariant. *Vision Research*, **32**, 465-475.
- STONER, G. R., ALBRIGHT, T. D., & RAMACHANDRAN, V. S. (1990). Transparency and coherence in human motion perception. *Nature*, **344**, 153-155.
- TURANO, K. (1991). Evidence for a common motion mechanism of luminance- and contrast-modulated patterns: Selective adaptation. *Perception*, **20**, 455-466.
- TURANO, K., & PANTLE, A. (1989). On the mechanism that encodes the movement of contrast variations: Velocity discrimination. *Vision Research*, **29**, 207-221.
- VAN SANTEN, J. P. H., & SPERLING, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America*, **A2**, 300-321.
- VICTOR, J., & CONTE, M. (1991). Coherence of Fourier and non-Fourier gratings: Cues of a feather flock together. *Investigative Ophthalmology & Visual Science*, **32**(Suppl.), 893.
- VICTOR, J. D., & CONTE, M. M. (1992). Evoked potential and psychophysical analysis of Fourier and non-Fourier motion mechanisms. *Visual Neuroscience*, **9**, 105-123.
- WILSON, H. R. (1991). A psychophysically motivated model for two-dimensional motion perception. *Investigative Ophthalmology & Visual Science*, **32**(Suppl.), 893.

(Manuscript received October 28, 1991;
revision accepted for publication March 30, 1992.)