

Motion Coherency Rules are Form-cue Invariant

GENE R. STONER,* THOMAS D. ALBRIGHT*†

Received 29 January 1991; in revised form 12 August 1991

Object features can be made manifest by differences in a variety of physical attributes or figural cues. In this study we provide evidence that motion signals arising from different figural cues can be combined to produce a percept of a coherently moving (heterogeneous-cue) pattern. Moreover, as is true for homogeneous-cue patterns, similarity along the dimensions of contrast (we introduce the idea of contrast equivalence) and spatial frequency determines whether coherent motion is perceived. These results are in accordance with recent physiological evidence demonstrating that directional selectivity of many neurons in cortical visual area MT is invariant over changes in the figural cue that defines the moving stimulus ("form-cue invariance").

Motion perception Plaid patterns Motion coherence Flicker Visual cortex Area MT

INTRODUCTION

Object features in our visual environment can be defined by a variety of physical attributes, including intensity and wavelength of reflected light, surface texture, relative motion, and distance from an observer. Our visual system readily constructs featural representations using many of the available cues. While information about cue identity is clearly retained and employed, our qualitative perceptual experience suggests that at some level featural representations may be "form-cue invariant". More specifically, we hypothesize the existence of featural representations that are uniformly sensitive to a broad range of cues for figure/ground segregation. It is thus appropriate to ask whether or not there are some essential functions performed by our visual system that operate upon purely cue invariant form information.

Motion processing is an obvious candidate. Psychophysical experiments have shown that motion can be perceived using stimuli defined by any of a variety of cues (e.g. Ramachandran, Rao & Vidyasagar, 1973a; Sperling, 1976; Julesz & Payne, 1968) suggesting the existence of a neural representation of motion that is insensitive to figural cue. Our recent physiological experiments addressing this hypothesis have established the existence of a population of directionally selective neurons in the primate middle temporal visual area (MT) that individually exhibit insensitivity to figural cue (Albright, 1987, 1992), thus mirroring the form-cue invariance manifested perceptually. The presence of such invariance at the single-neuron level leads to a number

of predictions. We will be concerned here with the possibility that human motion perception is insensitive to figural cue heterogeneity in moving patterns.

It has been proposed that detection of motion of two-dimensional patterns ("pattern-motion") involves integration of initial motion measurements that are local and one-dimensional. Considerable psychophysical (Adelson & Movshon, 1982; Movshon, Adelson, Gizzi & Newsome, 1985; Adelson, 1984; Welch, 1989; Krauskopf & Farell, 1990; Stoner, Albright & Ramachandran, 1990) and physiological (Gizzi *et al.*, 1983; Albright, 1984; Movshon *et al.*, 1985; Rodman & Albright, 1989; Stoner & Albright, 1991) evidence has accumulated in support of this possibility. Using a psychophysical paradigm and simple patterned stimuli ("plaids") developed by Adelson and Movshon (1982), this integration process has become amenable to study. A number of key parameters that affect motion signal integration have now been identified psychophysically. These parameters include: relative contrast and spatial frequency (Movshon *et al.*, 1985), perceptual transparency (Stoner *et al.*, 1990), color (Krauskopf & Farell, 1990; Kooi, DeValois, Grosf & Switkes, 1989) and binocular disparity (Adelson, 1984). Physiological experiments have revealed a sub-population of MT neurons that correctly encode pattern-motion (Movshon *et al.*, 1985; Rodman & Albright, 1989).

Consider a moving plaid pattern constructed such that its constituent moving gratings are defined by different figural cues—a "heterogeneous-cue" plaid pattern. It seems likely that some figural cue generalization will have occurred prior to the stage where the local motion signals are integrated. Indeed, the presence of significant populations of *both* pattern-motion neurons (25–30%)

*The Salk Institute for Biological Studies, La Jolla, CA 92186, U.S.A.
†To whom all correspondence should be addressed.

(Movshon *et al.*, 1985; Rodman & Albright, 1989) and form-cue invariant direction selectivity (50%) (Albright, 1987, 1992) in area MT makes a strong case that this system will combine motion signals from different figural cues to yield a coherent pattern motion percept. Nonetheless, critical physiological experiments have yet to be performed: pattern-motion selectivity and form-cue invariance have not been studied in the same experiments. The issue can, however, be addressed psychophysically. If the integration stage is form-cue invariant, we should expect this process to be insensitive to figural cue heterogeneity among the component motion signals. We have tested this prediction using plaid patterns designed such that each of the component gratings is defined by a different figure/ground cue. For these experiments, one component was always defined by conventional luminance contrast and the other by "flicker contrast", or relative temporal modulation. (Refer to Methods for details of this stimulus configuration.)

The specific goals of the present experiments were (1) to determine whether motion coherence could be perceived using components defined by different figural cues and, if so, (2) to examine the extent to which the integration process for these heterogeneous-cue patterns follows the same component similarity "rules" established for homogeneous-cue patterns (Adelson & Movshon, 1982). The latter should tell us something about form-cue invariance at the component integration stage.

In accordance with our prediction, human subjects experienced high rates of perceptual motion coherence when presented with "heterogeneous-cue" plaids constructed from luminance- and flicker-defined components. Moreover, the likelihood of such coherence exhibits a degree of dependence upon relative contrast and spatial frequency of the component gratings that is similar to what has been previously reported for conventional "homogeneous-cue" plaids.

METHODS

Visual stimuli

Visual stimuli consisted of moving "plaid" patterns produced by superposition of two drifting sinusoidally modulated gratings. One grating, hereafter referred to as the *luminance grating*, was constructed by sinusoidal modulation of the mean luminance of a randomly textured pattern [Fig. 1(B)]. Movement of this grating was achieved by drifting the sinusoid while the textured pattern remained stationary. In generating the luminance grating our aim was to reproduce as nearly as possible the conditions used in previous studies of motion coherence (e.g. Adelson & Movshon, 1982). The textured background was, nonetheless, necessitated by introduction of the flicker cue for motion (see below). While it is notable that the luminance grating alone contains competing cues at a single spatial location (stationary texture elements vs moving luminance modulation), motion perception is apparently unaffected for

the purposes of these experiments. Indeed, the luminance grating has the (quite natural) appearance of a shadow drifting smoothly across a textured field.

Superimposed upon the luminance cue was a second grating defined by sinusoidal modulation of the probability that individual texture elements will undergo contrast reversal [Fig. 1(C)]. Movement of this *flicker grating* was caused by a traveling wave of such (non-restorative) probabilistic contrast reversal, which gave the appearance of smoothly drifting bars composed of "twinkling" or "flickering" dots. This type of stimulus is a form of "motion without correlation", as originally described by Sperling (1976) and subsequently referred to by various authors (and in various incarnations) as moving "kinetic edges" (Anstis, 1980), "dynamic texture" (Anderson, Burt & van der Wal, 1985), " μ -motion" (Lelkens & Koenderink, 1984), "motion contrast" (Regan & Beverley, 1984) and "motion of successively generated subjective figures" (Petersik, Hicks & Pantle, 1978).

Of significance for the present experiments is the fact that the space-averaged luminance of the flicker grating is constant across all portions (moving foreground and static background) of the stimulus. Thus there are no coherent luminance gradients that define the moving flicker grating. Rather, it is detectable solely by virtue of differences in temporal frequency or flicker. The absence of luminance differences in the flicker grating was verified using a standard spot photometer (United Detector Technology, Hawthorne, Calif.). Due to stochastic fluctuations in the ratio of white to black texture elements, small local space-time variations in luminance occur. However, these random luminance variations lack spatiotemporal consistency and yield no information about the motion of the grating. Chubb and Sperling have applied the term "non-Fourier" to motion of this type since, in contrast to conventional luminance-based motion (such as the luminance grating used in these experiments), it cannot be uniquely characterized by its spatiotemporal Fourier power spectrum. More generally, this stimulus is a member of the class of "second-order" motion stimuli described by Cavanagh and Mather (1989), which also includes those defined by stereoscopic and spatial texture cues.

The luminance and flicker gratings represent orthogonal dimensions in a multidimensional *physical* (but not *perceptual*, we hypothesize) contrast space. When spatially superimposed to form a plaid pattern, modulation of each grating was thus entirely independent of the other. Throughout this report we refer loosely to these luminance/flicker plaid patterns as "heterogeneous-cue" plaids, which indeed they are. It should be recognized, however, that they represent only one of many potential combinations of figural cues and thus the generality of our findings remains a matter of speculation. This particular combination of cues was chosen in view of the fact that many directionally selective MT neurons have been shown to exhibit similar direction tuning for the same pair of cues (Albright, 1987, 1992).

A

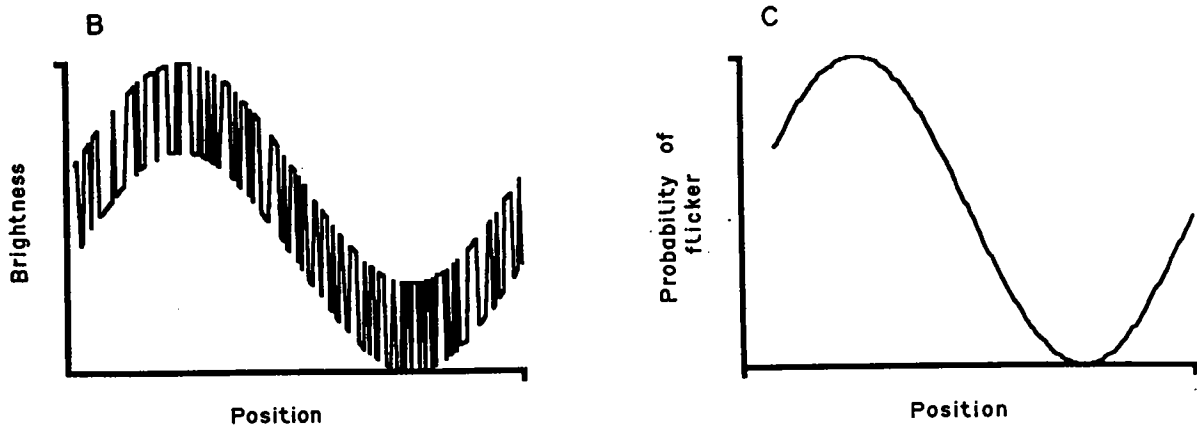
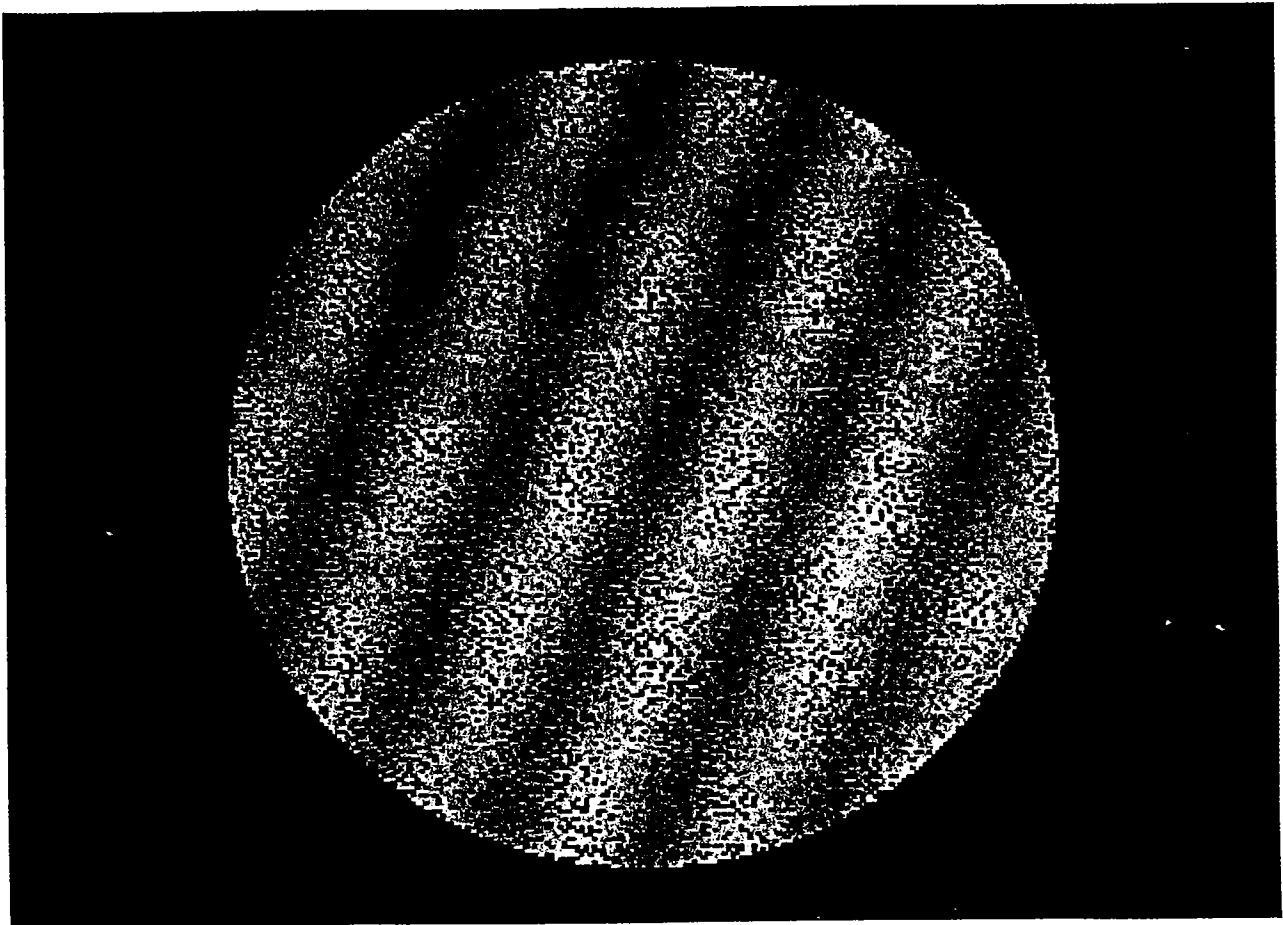


FIGURE 1. Schematic depiction of "heterogeneous-cue" plaid and procedure used to generate this stimulus. Both gratings were produced by appropriate (luminance or flicker) modulation of a randomly textured field. Component gratings were oriented at 45° relative to one another and were moved in directions differing by 135° . Resultant plaid pattern motion, when perceived, was either directly up or down, i.e. falling between the two component directions. See text for additional details regarding stimulus generation and viewing conditions. *Luminance contrast grating*: luminance contrast modulation is achieved by varying mean luminance of light and dark dots while keeping local dot difference constant. As shown in (B), mean intensity varies sinusoidally in a cross-section through this grating. *Flicker contrast grating*: colored dots in (A) have been used to indicate flicker contrast grating. (Actual dots are not colored.) The dots represent a set of randomly chosen elements that might undergo contrast reversal (i.e. flicker) at one instant in time. As shown in (C), probability of flicker varies sinusoidally in cross-section through this grating. Motion of the grating is achieved by a traveling wave of such flicker. We define flicker contrast amplitude in units of flicker probability.

Contrast (Michelson) of the luminance cue grating was defined as $(\max \text{ luminance} - \min \text{ luminance}) / (\max \text{ luminance} + \min \text{ luminance})$ (luminance here refers to the local average luminance of the textured pattern—see Fig. 1) and was varied according to the conditions of the experiment (see below), while flicker contrast, defined as $(\max \text{ flicker probability} - \min \text{ flicker probability}) / (\max \text{ flicker probability} + \min \text{ flicker probability})$ was held constant. Likewise, spatial frequency of the luminance grating was varied according to experimental conditions (see below), while spatial frequency of the flicker grating remained constant. The two gratings were oriented at 45° to one another. Each grating was moved in a direction perpendicular to its orientation; together the two gratings moved in directions differing by 135° . Each was drifted at $3^\circ/\text{sec}$. Predicted plaid direction was always intermediate between the two components (either up or down) with a speed of $7.75^\circ/\text{sec}$. Minimum dot size was 0.067° (1 pixel), aspect ratio was square, and dot density was 50%. The local dot luminance difference was constant (approx. 90 cd/m^2) and thus independent of the phase of the luminance grating [see Fig. 1(B)]. The space-averaged luminance of the stimulus display, with each luminance-modulated grating superimposed upon the textured field, was 56 cd/m^2 . During the inter-trial interval the textured field was visible with mean luminance equal to that of the stimuli. Room lighting was about 0.5 cd/m^2 and stimuli were viewed binocularly through natural pupils.

All stimuli were generated using a high-resolution graphics display controller (Pepper SGT, Number Nine Computer Corporation: 640×480 pixels, 8 bits/pixel, 60 Hz, non-interlaced) operating in an AT (80286) personal computer. Stimuli were displayed on a 14" analog RGB video monitor (Zenith ZCM-1490, flat technology CRT, phosphor P-22) and were viewed through a circular aperture subtending 11° at a distance of 43 cm from the nodal point of the eye. Photometric linearization tables were computed and used to reform the highly non-linear 8 bit voltage-luminance relationship. Movement was achieved by updating the position of each grating in synchrony with the vertical refresh of the display on alternate cycles (i.e. every 33.3 msec).

Psychophysical methods

The effects of relative contrast and spatial frequency on motion coherence for our heterogeneous-cue plaids were investigated in a two alternative forced-choice procedure using the method of constant stimuli. As in previous studies of motion coherence employing plaid patterns, subjects were instructed to report whether the heterogeneous-cue plaids moved as a single pattern or, alternatively, as two gratings sliding across one another. Judgments of this sort are criterion-dependent (Movshon *et al.*, 1985; Welch, 1989) and for this reason we have refrained from pooling data or making comparisons across subjects. Rather, we were largely concerned with general features of the variations in perceived motion coherence for individual subjects across the different stimulus conditions.

All subjects were first presented with a series of practice trials using conventional homogeneous-cue (luminance/luminance) plaids at a range of relative contrasts and spatial frequencies. Subjects were instructed to fixate a small cross at the center of the display for the duration of each 2.0 sec stimulus exposure and to indicate their dominant percept (component or pattern motion) with an appropriate key-press at the end of each trial. This practice continued until subjects expressed confidence in their perceptual judgments. Judgment confidence increases with the development of a stable internal criterion for making component vs pattern decisions. Subjects differed slightly in the number of trials needed to develop a stable criterion but in all cases practice continued for 20–30 trials. Subjects were then presented with a second series of practice trials using heterogeneous-cue plaids. These trials continued until subjects were again confident in the judgments—in all cases 20–30 trials. For heterogeneous-cue plaids the "reference" (non-variable parameters) grating was always flicker contrast while the "test" (variable parameters) grating was always luminance contrast.

Experiment 1 examined the effects of relative contrast on coherence of heterogeneous-cue plaids. Flicker contrast was fixed at one of three values (0.42, 0.58, or 0.83) while luminance contrast varied over a 3.5 octave range (0.017–0.200). Spatial frequency of both gratings was constant at 0.5 c/deg (a value determined to be optimal for coherence in preliminary testing). To ensure that gratings having low luminance contrast did not fail to cohere simply because they were below detection threshold, subjects were shown the same stimuli in a separate block of trials and were asked to report only the presence or absence of the luminance grating.

Experiment 2 examined the effects of relative spatial frequency on coherence. Flicker spatial frequency was fixed at one of three values (0.71, 0.5, or 0.35 c/deg) while luminance spatial frequency varied over a 3.75 octave range (0.18–2.50 c/deg) centered on that for the flicker grating. Flicker contrast was arbitrarily fixed at 0.83 as in Experiment 1 and luminance contrast was held at the value determined to be optimal for coherence at the spatial frequency used in Experiment 1.

Experiments 1 and 2 can be thought of as obtaining representative cross-sections through the surface that defines probability of coherence as a function of both relative contrast and spatial frequency. The effects of these variables, however, are not necessarily independent. For this reason we have also sparsely sampled the entire surface (8 contrast levels \times 7 spatial frequencies) in one subject (GS).

All of the stimulus conditions were presented in pseudo-random order for blocks of 8 trials/stimulus within each experiment. The data reported here represent mean coherence probability based upon 32 presentations of each stimulus.

Subjects

Two of the subjects (GT and PR) were inexperienced psychophysical observers with normal acuity and were

naive to the purposes of the experiment. Two additional subjects were the authors (GS and TA), both of whom were experienced psychophysical observers with corrected myopia.

RESULTS

All subjects reported coherent motion of heterogeneous-cue (luminance/flicker) plaid patterns under a limited range of stimulus conditions.

Experiment 1: effects of relative contrast

Adelson and Movshon (1982) observed maximal coherence of homogeneous-cue plaid patterns when component gratings were of identical luminance contrast. Extrapolating from this result in light of the form-cue invariance of many MT neurons, we predict maximal coherence of heterogeneous-cue plaids when the gratings are of equivalent *perceptual contrast* or *saliency*. At present we have no independent measure of the percep-

tual saliency of the cues used in this experiment. If our hypothesis is correct, however, we should expect to find a single peak in the relative contrast tuning function. Furthermore, the luminance contrast value yielding peak coherence should vary in accordance with changes in the flicker contrast value used for the reference grating.

The effects of relative contrast are illustrated in Fig. 2. Each of our subjects reported coherence of heterogeneous-cue plaid patterns over a narrow contrast range. Moreover, with one exception (subject PR) the probability of coherence could be accurately described by a Gaussian function of log luminance contrast. Using a flicker grating contrast of 0.83 the location of the peak of this Gaussian was 0.090, 0.073, and 0.077 respectively for subjects GS, TA, and GT. This is the luminance contrast value we interpret to be of equivalent perceptual contrast (for each subject) to this "standard" flicker grating used in these experiments. Indeed, each of these subjects reported nearly 100% motion coherence at luminance values near this point. Coherence probability

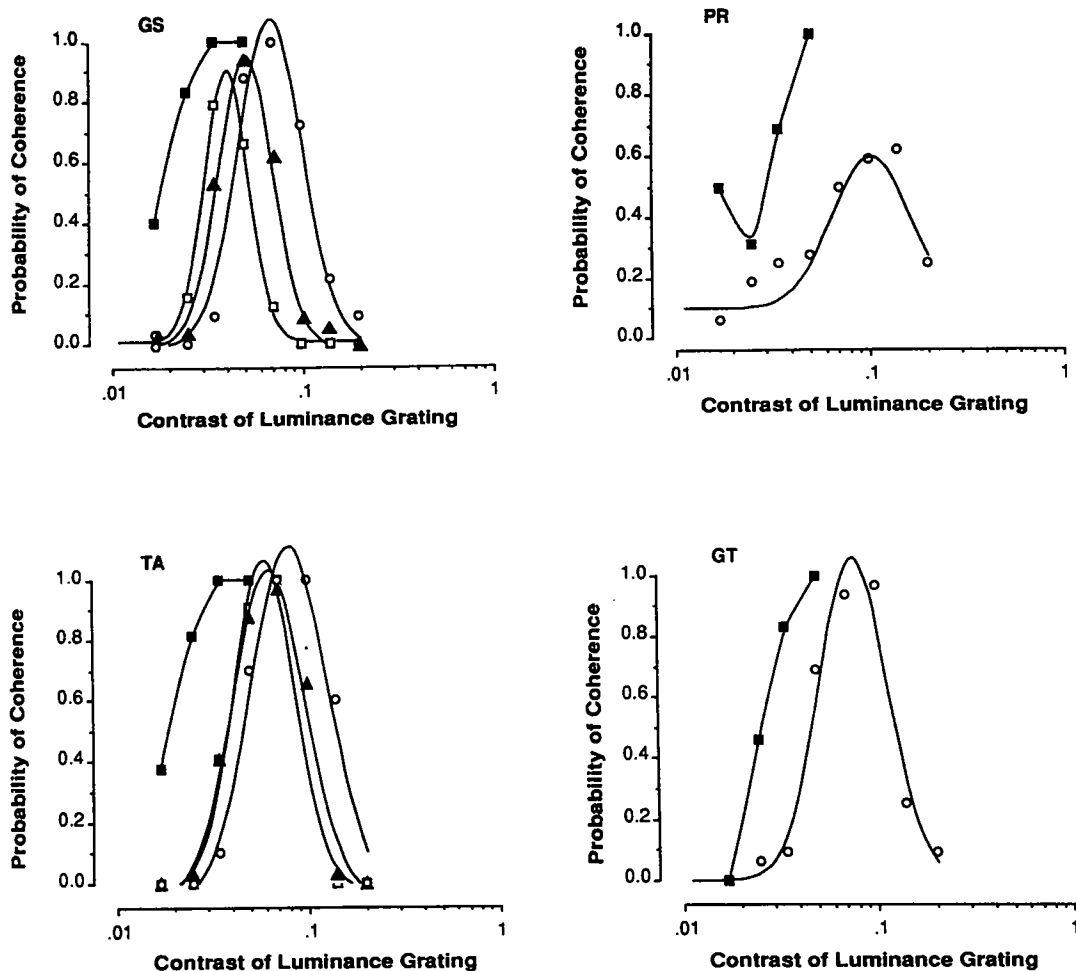


FIGURE 2. Effects of relative contrast on motion coherence for heterogeneous-cue gratings. The contrast of the flicker contrast (reference) grating was held constant across trials while luminance grating (test) varied from 0.017 to 0.200 contrast. Probability of motion coherence is plotted as a function of test contrast. All subjects experienced maximal coherence over a narrow contrast range. All four subjects were studied using the "standard" flicker contrast (reference) value of 0.83 (minimum probability of flicker = 0.02, maximum = 0.22). These data are identified by the open circles. Two subjects (GS and TA) were also studied using smaller flicker contrast values of 0.58 (0.05 minimum, 0.19 maximum) and 0.42 (0.07 minimum, 0.17 maximum), identified by solid triangles and open squares, respectively. Note that these smaller reference contrasts are associated with leftward shifts of the motion coherence functions. Luminance contrast detection curves are also shown (solid squares) for each subject (for these curves the y -axis shows probability of *detection* rather than coherence). Each point represents 32 trials.

dropped off sharply with deviations from the optimal luminance contrast; full-width, half-maximum tuning bandwidths (calculated from the standard deviation of the fitted Gaussian) were 2.19, 1.72, and 1.37 octaves for subjects GS, TA, and GT. Data for subject PR was, in general, more variable with significantly lower coherence rates. Coherence probability for this subject was, nonetheless, a unimodal function of relative contrast with a peak at 0.103 luminance contrast—a value slightly higher than that for any of the other subjects.

Luminance contrast detection curves obtained from each subject are also plotted in Fig. 2. In all cases coherence dropped to near zero probability at luminance contrast levels above detection threshold. These results demonstrate that the decline in pattern motion percepts cannot be attributed to a simple failure to detect one of the components. This conclusion is, moreover, supported by the fact that coherence rates decline with *increasing* as well as decreasing luminance contrast levels.

Our assertion that maximal coherence occurs at the level of equivalent contrast or salience leads us to predict shifts in the optimal luminance (test) contrast that follow

shifts in the flicker (reference) contrast, so that an increase or decrease in the flicker (reference) grating contrast would lead to a corresponding increase or decrease in the optimal luminance (test) contrast. Two subjects (GS and TA) were tested with flicker contrast values that were both lower (0.42 and 0.58) than the standard (0.83). These data are also shown in Fig. 2. Once again, high coherence rates were reported over a narrow contrast range. Moreover, the midpoint of this range shifted leftward in accord with our predictions.

Experiment 2: effects of relative spatial frequency

Adelson and Movshon (1982) observed maximal coherence of homogeneous-cue plaid patterns when component gratings were of identical spatial frequency. Our hypothesis of form-cue insensitivity at the motion integration stage predicts a similar dependence upon relative spatial frequency for heterogeneous-cue plaids.

The effects of relative spatial frequency are illustrated in Fig. 3. Two subjects (PR and TA) were tested using a single reference frequency (0.5 c/deg), one subject (GT) was tested using two different reference frequencies (0.71 and 0.35 c/deg) and a fourth subject (GS) was tested

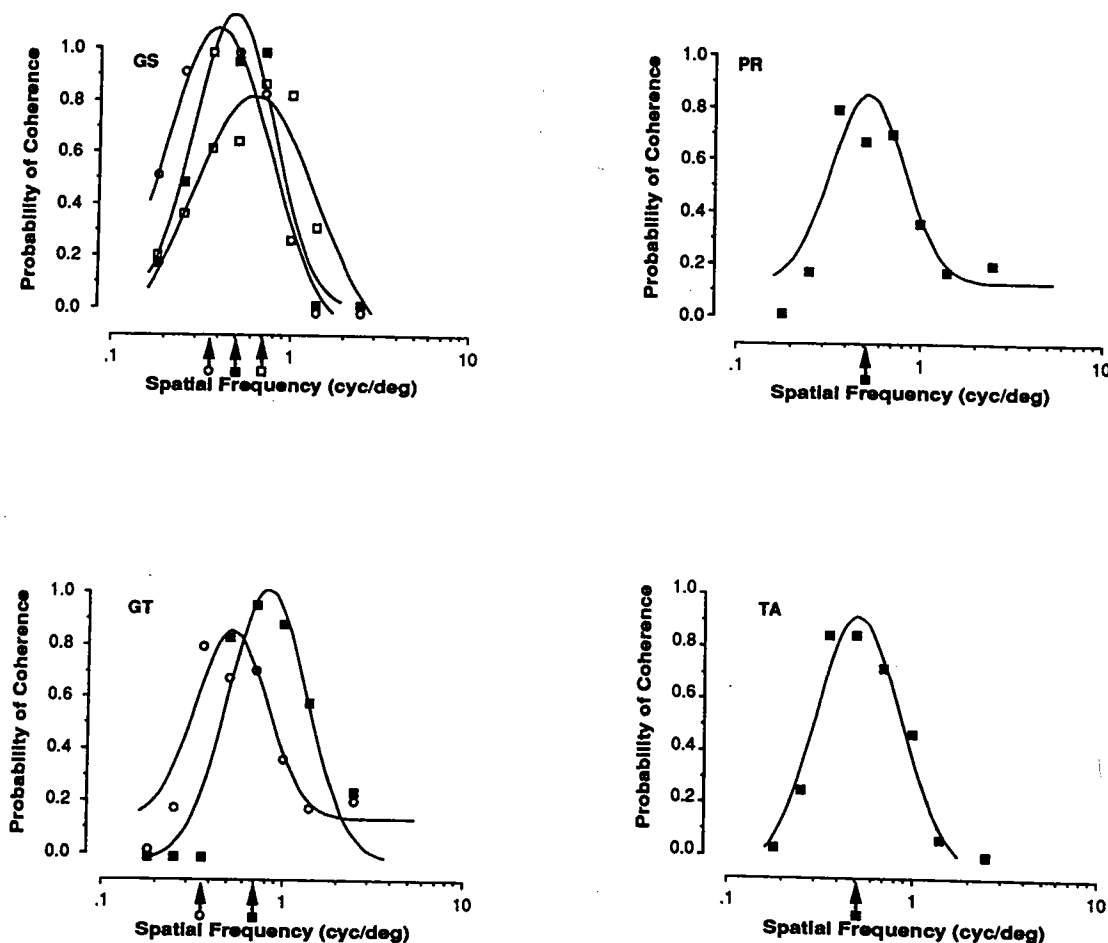


FIGURE 3. Effects of relative spatial frequency on motion coherence for heterogeneous-cue plaids. Frequency and contrast of flicker grating (reference) were constant (0.35 or 0.5 or 0.71 c/deg, and 0.83 flicker contrast) while frequency of luminance grating (test) was varied from 0.18 to 2.50 c/deg. Luminance contrast was constant. Probability of coherence is plotted as a function of test spatial frequency. All subjects experienced maximal coherence over a narrow frequency range centered at reference (flicker contrast) grating frequency. Spatial frequency of each reference grating is indicated by arrows along abscissa. Each point represents 32 trials.

using three reference frequencies (0.35, 0.50, and 0.71 c/deg). In every case subjects reported high rates of coherence (65–100%) of heterogeneous-cue plaid patterns over a narrow range of test spatial frequencies centered around the frequency of the reference grating.

The relative frequency tuning curves were each fitted with a Gaussian function of log spatial frequency of the luminance grating. As shown in Fig. 3, the locations of these Gaussian peaks correspond well to the reference frequencies used. The full-width, half-maximum tuning

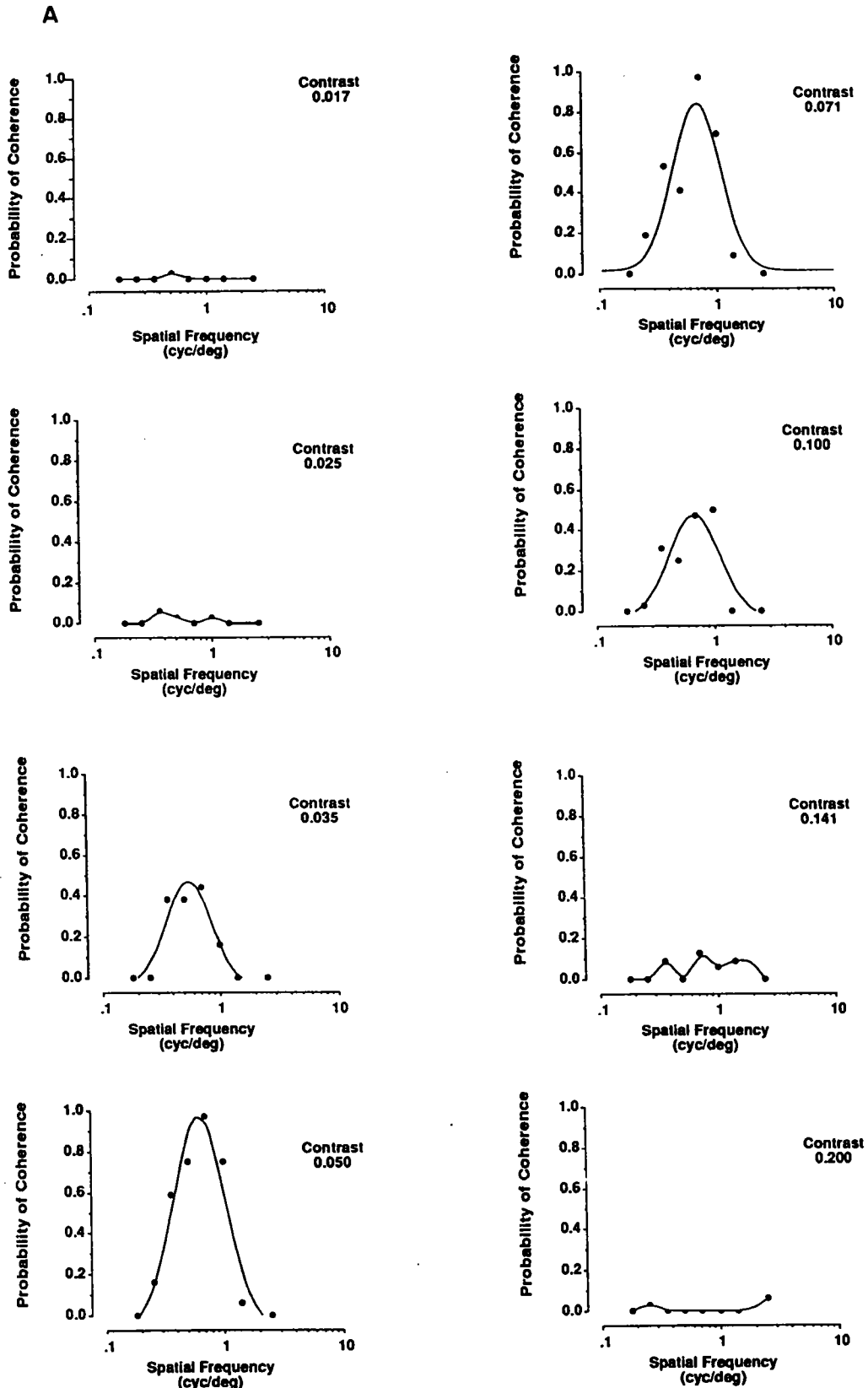


FIGURE 4. (A) *Caption overleaf.*

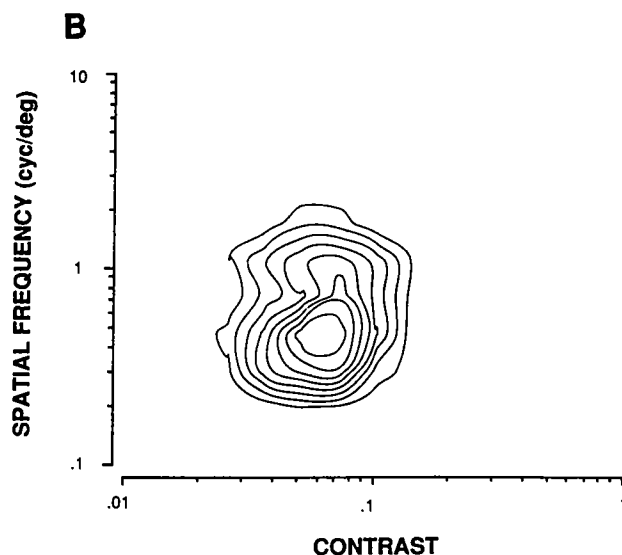


FIGURE 4. Effects of relative contrast and spatial frequency of heterogeneous-cue gratings on motion coherence for subject GS. (A) Probability of coherence as a function of test grating (luminance contrast) spatial frequency for eight different luminance contrast levels. The reference grating (flicker contrast) spatial frequency (0.5 c/deg) and contrast (0.83 flicker) were fixed. Each data point represents 32 trials. High rates of perceptual coherence were seen only for intermediate luminance contrast levels of the test grating (0.035, 0.050, 0.071 and 0.100) and over a narrow range of test grating spatial frequencies. Maximal coherence was thus reported over a small region in contrast-frequency space. (B) Contour plot of coherence probability in contrast-frequency space for the data shown in (A). Iso-probability lines were interpolated at increments of 0.1 probability. The peak in this function may identify the point at which the two gratings are of equal perceptual salience.

bandwidth was 1.89 octaves, on average, and was also quite consistent across reference conditions and subjects (range, 1.37–2.75 octaves).

Contrast/spatial frequency response surface

The relative contrast and spatial frequency tuning curves presented above are merely cross-sections through the surface defined by probability of coherence as a function of these two variables. We had little reason to believe that contrast and spatial frequency have any interactive effects on probability of coherence. We, therefore, predicted that the peaks of the Gaussians fitted to probability of coherence as a function of spatial frequency would remain relatively constant across changes in contrast. To investigate this possibility we tested one subject (GS) with a full complement of test (luminance) contrast and spatial frequency gratings (8 contrast \times 7 frequency). These test grating parameters were centered around the reference (flicker) grating, which was fixed at 0.83 contrast and 0.5 c/deg.

The spatial frequency tuning curves are shown for each of the 8 test contrast levels in Fig. 4(A). Coherence probability never reached above 20% for very low (0.017 and 0.025) or very high (0.141 and 0.200) reference contrasts. Each of the remaining curves (contrast 0.035, 0.050, 0.071, and 0.100) was fitted with a Gaussian, which revealed peak coherence probability at spatial frequencies of 0.55, 0.63, 0.75, and 0.69, respectively. Full-width, half-maximum tuning bandwidths was also very consistent across these different contrast levels (1.53, 1.73, 1.95, and 1.68 octaves). These results indicate that, where motion coherence does occur for heterogeneous-cue plaids, the coherence exhibits approxi-

mately the same relative spatial frequency dependence across a range of relative contrast levels.

Figure 4(B) contains a contour plot of these coherence probability data in contrast/spatial frequency space. Maximal coherence was reported only over a small region in this space. Moreover, the decline in coherence was roughly symmetric about the peak, providing little evidence for any interactive effects of contrast and spatial frequency. The similarity of reference and test frequency is consistent with the results of Experiment 2 and the peak contrast level is believed to identify the point at which the two cues are of equal perceptual salience.

DISCUSSION

Real object motion can result in a variety of different space-time patterns of luminance. Consistent with this physical reality is the qualitative similarity of motion percepts elicited by different figural cues. The present results demonstrate that motion signals arising from two different figural cues (luminance and flicker) can elicit quantitatively similar effects at the level of motion signal integration. They demonstrate that the perception of coherent motion is dependent upon component similarity in the dimensions of spatial frequency and contrast even when these components are defined by cues that are physically very different. These data reveal that the contrast and spatial frequency mechanisms, which gate the coherence process, are form-cue invariant—at least for luminance and flicker contrast cues.

The present results thus extend the concept of form-cue invariance, originally applied to first-stage motion

detection by MT neurons (Albright, 1987, 1992), to include the second-stage motion signal integration ("pattern motion") mechanism that is presumed to reside in area MT. It is important to emphasize once again, however, that our experiments have tested motion coherence using only a limited subset of the possible figural cues, namely luminance and flicker. It would, therefore, be premature to conclude from our results that this coherence mechanism possesses more than a "limited" form-cue invariance. Nevertheless, given the computational appeal of form-cue invariance, the adaptive value that might be afforded by a truly form-cue invariant system, the aforementioned neurophysiological data and additional psychophysical evidence bearing upon the issue (see below), it is tempting to generalize from the present results and speculate that the motion coherence mechanism in MT may be broadly insensitive to figural cue.

In the remainder of this discussion we consider (1) the neural origin of form-cue invariance, (2) the implications for models of motion signal integration, (3) the notion of "contrast equivalence" for stimuli defined by different cues, (4) the relevance to other psychophysical studies using motion stimuli defined by different figural cues, and (5) the significance of form-cue invariant motion processing.

Neural origin of form-cue invariance

The diversity of figural cues that can define features requires a diversity of mechanisms for their detection. Features defined by motion, color, disparity or texture each require a different type of detector and all are invisible to conventional luminance based detectors (see Chubb & Sperling, 1988). However, on grounds of computational economy it would be advantageous for features defined by different cues to be encoded by the same neurons when the identity of the form-cue is not critical for the computation at hand. Physiological evidence that the motion system discards form-cue identity was first demonstrated by Albright (1987). Many cells in area MT of the macaque were shown to exhibit similar direction tuning for moving bars defined by texture, luminance or flicker contrast. In addition, some MT neurons are directionally selective for stimuli defined solely by chromatic contrast (Saito, Tanaka, Isono, Yasuda & Mikami, 1989; Charles & Logothetis, 1989; Dobkins & Albright, 1991). Area MT is not the only region of extrastriate visual cortex where evidence for cue generalization has been observed: Logothetis and Charles (1990) have reported that cells in area V4 exhibit orientation tuning for flicker-defined stimuli, which is similar to that observed for luminance-defined stimuli. But where might form-cue invariance first be expressed?

Albright and Chaudhuri (1989) have recently discovered that many neurons in area VI of the macaque exhibit similar orientation and direction tuning for stimuli defined either by flicker contrast or luminance contrast. There is also evidence that some degree of form-cue invariance begins as early as the LGN. Derington and Lennie (1984) reported that magnocellular neurons of the LGN exhibit different responsivities to

temporal modulation in their surround and center. This result suggests that magnocellular LGN neurons might respond to stimuli defined by flicker contrast, in addition to their well-documented luminance contrast sensitivity.

Of course, some types of form-cues require more processing steps to extract and could not be expected to occur in the LGN. For example, borders defined by stereoscopic disparity could not be detected prior to binocular convergence and "subjective contours" require preliminary extraction of "real" contours and are not found until area V2 (von der Heydt & Peterhans, 1989; Peterhans & von der Heydt, 1989).

Relevance to models of motion signal integration

The perceived direction of plaid stimuli could result from two distinct mechanisms: (1) the velocities of the oriented component gratings could first be extracted and then combined to compute pattern direction; and (2) the "blobs" formed at the intersections could be tracked. There is strong physiological and psychophysical data in support of the first alternative (Movshon *et al.*, 1985; Welch, 1989). It has been suggested (Cavanagh & Mather, 1989) that the coherence of heterogeneous-cue or "multi-attribute" plaids might provide additional evidence against the blob tracking alternative because the intersections so formed would seem to require a complicated kind of detector. However, we see no *a priori* reason to believe that heterogeneous-cue blob detectors are more "complicated" or less likely to exist than heterogeneous-cue integrators of one-dimensional motion. In fact, we would predict that magnocellular LGN neurons would respond to intersections like those found in the stimuli used in these experiments (though not to their motion). Therefore, we believe that our results must be considered neutral on this question.

Contrast equivalence

For two moving luminance gratings, maximal coherence is obtained when they are of the same spatial frequency and contrast (Adelson & Movshon, 1982), suggesting that the motion integration process is gated by contrast and spatial frequency mechanisms (expressed as either labeled line or firing rate coding). Our demonstration that stimuli with very different physical characteristics can activate these processes in very much the same way suggests that they are causing equivalent types of activation in the relevant mechanisms. They are thus, for all practical purposes, of equivalent contrast and spatial frequency, despite the obvious differences in the patterns of luminance.

Spatial frequency maps quite naturally across figural cues, but contrast requires an arbitrary mapping. This prompts us to introduce the concept of "contrast equivalence". Contrast equivalence identified in our coherence paradigm might be expected to extend to other psychophysical tasks. For example, contrast equivalent stimuli might have similar detection thresholds. Physiologically, we would expect stimuli of equivalent contrast, so defined, to cause a similar level of activation of motion sensitive neurons that exhibit form-cue invariance.

Relevance to other studies

Cavanagh, Arguin, and von Grunau (1989) compared apparent motion percepts defined by intra-attribute stimuli with that of inter-attribute stimuli. The stimulus consisted of two disks alternating at 2.0 Hz. They reported that the "motion strength" for inter-attribute stimuli (including luminance, color, binocular disparity, texture, and motion defined disks) was about 75% of that measured for intra-attribute stimuli. This is consistent with the findings reported here as well as with the physiological demonstration that form-cue invariance extends only to a subset of motion sensitive neurons (Albright, 1987, 1992).

Cavanagh and Mather (1989) considered the percepts resulting from superimposing two gratings of the same orientation, but moving in opposite directions. They compared the case in which each grating is defined by the same cue (dynamic texture or luminance contrast) with the case in which each grating is defined by a different cue. Two different percepts can occur under these stimulus conditions: (1) *competition*, in which motion is perceived in only one direction or the other; and (2) *transparency* in which the opposing motion of both gratings is clearly perceived. They found that the results were the same in each case: competition when the gratings were of different spatial frequencies and transparency when they were of similar spatial frequencies. They also reported that when the different gratings were of the same spatial frequency, competition occurred when the luminance grating was adjusted to a particular contrast. These results are remarkably consistent with our findings, including the implication of contrast equivalence.

Significance of form-cue invariant motion signal integration

We have shown that the motion system avoids computational redundancy by dispensing with information about form-cue identity. There is evidence that the visual system does this for other visual tasks as well. For example, Cavanagh (1987) found evidence for "featural integration" in perception of perspective. The evidence for form-cue invariance in stereopsis is mixed. Cavanagh (1987) found that no depth was perceived when different form-cues were used in each eye, whereas Ramachandran *et al.* (1973b) reported that a luminance border from one eye can be fused with a disparate chromatic or texture border from the other eye.

It is clear that for some perceptual tasks, the identity of the form-cue is important. The perception of shape from shading and surface relief defined by shadows also disappears when the stimulus is defined by cues other than luminance (Cavanagh & Leclerc, 1989). This lack of perceptual form-cue invariance is consistent with the fact that the information provided by luminance in these examples is specific to luminance. Perceiving cues other than luminance as generating shape from shading or shadows is simply not consistent with the real world and form-cue invariance in those cases would generate spurious percepts. Computational economy has apparently

been weighed against veridicality in determining which perceptual tasks can operate on generalized featural information.

Computational economy, however, is not the only advantage of form-cue invariant directional selectivity. The motion of real objects will usually be defined by several figural cues. For instance, the motion of a predator moving through dense brush might be detected via motion contrast boundaries created by the motion of the brush, or by glimpsing luminance, chromatic, texture or depth boundaries. Locally one cue will almost certainly be more salient than another. It would appear advantageous to be able to rapidly integrate information from these different cues in order to compute true object velocity. This is a property which emerges naturally from a system which discards figure cue identity.

REFERENCES

- Adelson, E. H. (1984). Binocular disparity and the computation of two-dimensional motion. *Journal of the Optical Society of America*, *A1*, 1266.
- Adelson, E. H. & Movshon, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, *300*, 523-525.
- Albright, T. D. (1987). Isoluminant motion processing in macaque visual area MT. *Society of Neuroscience Abstracts*, *13*, 1626.
- Albright, T. D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. *Journal of Neurophysiology*, *52*, 1106-1130.
- Albright, T. D. (1992). Form-cue invariant motion processing in primate visual cortex. *Science*. Submitted.
- Albright, T. D. & Chaudhuri, A. (1989). Orientation selective responses to motion contrast boundaries in macaque VI. *Society of Neuroscience Abstracts*, *15*, 323.
- Anderson, C. H., Burt, P. J. & van der Wal, G. S. (1985). Change detection and tracking using pyramid transform techniques. *SPIE Intelligent Robots and Computer Vision*, *579*, 71-78.
- Anstis, S. M. (1980). The perception of apparent movement. *Philosophical Transactions of the Royal Society of London B*, *290*, 153-168.
- Cavanagh, P. (1987). Reconstructing the third dimension: Interactions between color, texture, motion, binocular disparity, and shape. *Computer Vision, Graphics, and Image Processing*, *37*, 171-195.
- Cavanagh, P. & Leclerc, Y. G. (1989). Shape from shadows. *Journal of Experimental Psychology: Human Perception and Performance*, *15*, 3-27.
- Cavanagh, P. & Mather, G. (1989). Motion: The long and short of it. *Spatial Vision*, *4*, 103-109.
- Cavanagh, P., Arguin, M. & von Grunau, M. (1989). Interattribute apparent motion. *Vision Research*, *29*, 1197-1204.
- Charles, E. R. & Logothetis, N. K. (1989). The responses of middle temporal (MT) neurons to isoluminant stimuli. *Investigative Ophthalmology and Visual Science*, *30*, 427.
- Chubb, C. & Sperling, G. (1988). Drift-balanced stimuli: A general basis for studying non-Fourier motion perception. *Journal of the Optical Society of America*, *A5*, 1986-2006.
- Derrington, A. M. & Lennie, P. (1984). Spatial and temporal contrast sensitivities of neurones in lateral geniculate nucleus of macaque. *Journal of Physiology*, *357*, 219-240.
- Dobkins, K. R. & Albright, R. D. (1991). What happens if it changes color when it moves? *Investigative Ophthalmology and Visual Science*, *32*, 822.
- Gizzi, M. S., Newsome, W. T. & Movshon, J. A. (1983). Directional selectivity of neurons in macaque MT. *Investigative Ophthalmology and Visual Science (Suppl.)*, *24*, 107.
- Gorea, A. & Lorenceau, J. (1989). Motion perception in compound stimuli is "blob"-dependent. *Investigative Ophthalmology and Visual Science (Suppl.)*, *30*, 388.

- Julesz, B. & Payne, R. A. (1968). Difference between monocular and binocular stroboscopic movement perception. *Vision Research*, 8, 433-444.
- Kooi, F. L., DeValois, K. K., Grosf, D. H. & Switkes, E. (1989). Coherence properties of colored moving plaids. *Investigative Ophthalmology and Visual Science (Suppl.)*, 30, 389.
- Krauskopf, J. & Farell, B. (1990). Influence of color on the perception of coherent motion. *Nature*, 348, 328-331.
- Lelkens, A. M. M. & Koenderink, J. J. (1984). Illusory motion in visual displays. *Vision Research*, 24, 1083-1090.
- Logothetis, N. K. & Charles, E. R. (1990). V4 responses to gratings defined by random dot motion. *Investigative Ophthalmology and Visual Science*, 31, 444.
- Movshon, J. A., Adelson, E. A., Gizzi, M. & Newsome, W. T. (1985). The analysis of moving visual patterns. In Chagas, C., Gattass, R. & Cross, C. G. (Eds), *Study group on pattern recognition mechanisms*. Vatican City: Pontifica Academia Scientiarum.
- Peterhans, E. & von der Heydt, R. (1989). Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. *Journal of Neuroscience*, 9, 1749-1763.
- Petersik, J. T., Hicks, K. I. & Pantle, A. J. (1978). Apparent movement of successively generated subjective figures. *Perception*, 7, 371-383.
- Ramachandran, V. S., Rao, V. M. & Vidyasagar, T. R. (1973a). Apparent motion with subjective contours. *Vision Research*, 26, 1969-1975.
- Ramachandran, V. S., Rao, V. M. & Vidyasagar, T. R. (1973b). The role of contours in stereopsis. *Nature*, 242, 412-414.
- Regan, D. & Beverley, K. I. (1984). Figure ground segregation by motion contrast and by luminance contrast. *Journal of the Optical Society of America*, A1, 433-442.
- Rodman, H. R. & Albright, T. D. (1989). Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Experimental Brain Research*, 75, 53-64.
- Saito, H., Tanaka, K., Isono, H., Yasuda, M. & Mikami, A. (1989). Directionally selective responses of cells in the middle temporal area (MT) of the macaque monkey to the movement of equiluminous opponent color stimuli. *Experimental Brain Research*, 75, 1-14.
- Sperling, G. (1976). Movement perception in computer-driven visual displays. *Behavioral Research Methods and Instrumentation*, 8, 144-151.
- Stoner, G. R. & Albright, T. D. (1991). Responses of area MT neurons to non-coherently moving plaid patterns. *Investigative Ophthalmology and Visual Science*, 32, 822.
- Stoner, G. R., Albright, T. D. & Ramachandran, V. S. (1990). Transparency and coherence in human motion perception. *Nature*, 344, 153-155.
- Von der Heydt, R. & Peterhans, E. (1989). Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. *Journal of Neuroscience*, 9, 1731-1748.
- Welch, L. (1989). The perception of moving plaids reveals two motion-processing stages. *Nature*, 337, 734-736.

Acknowledgements—We thank K. Dobkins, A. Chaudhuri, F. Nahm, S. LeVay, T. Sejnowski, and G. Carman for helpful discussions and comments on the manuscript, and J. Costanza for superb technical assistance. This work was supported by National Eye Institute Grant EY-07605, an Alfred P. Sloan Foundation Research Fellowship, and a McKnight Foundation Neuroscience Research Development Award (TDA). G. R. Stoner was partially supported by a fellowship from the McDonnell-Pew Center for Cognitive Neuroscience at San Diego.

