Detection and discrimination of coherent motion

LYN MOWAFY, RANDOLPH BLAKE, and JOSEPH S. LAPPIN
Vanderbilt University, Nashville, Tennessee

When viewing a pair of bars defined by the difference of spatial Gaussian functions (DOGs), human observers can discriminate accurately the relative movements of the bars, even when they differ in spatial frequency. On each trial, observers viewed two brief presentation intervals in which a pair of vertically oriented DOGs moved randomly back and forth within a restricted range. During one interval, both bars moved in the same horizontal direction and by the same magnitude (correlated movements); in the other interval, their movements were uncorrelated. When discrimination accuracy is related to the simultaneous detection of two independent movements, it was found that, if observers can detect the movements of spatially separated bars, they can tell whether their relative movements are correlated. Performance remained remarkably accurate even when the two bars differed in spatial frequency by more than two octaves or were presented separately to the two eyes. Apparently, the accurate discrimination of coherent motion involves an efficient spatial integration of optical motion information over multiple spatial locations and multiple spatial scales.

The human visual system is remarkably adept at detecting motion: contour displacements of just a few arc seconds may be detected under optimal conditions (Legge & Campbell, 1981). Suppose, however, that two moving contours are spatially separated by several degrees and that observers are asked whether the concurrent displacements of these contours have the same directions and distances. How accurately can observers discriminate common or "coherent" movements from uncorrelated movements? Moreover, how is the discriminability of coherent motion affected by differences in the spatial frequency content of the two moving contours?

The latter question is of interest because spatially separate components of a given object may be detected by mechanisms with different scales of spatial resolution when they are defined by different levels of structural detail or are separated in depth by distances that impose different degrees of accommodative resolution. Regardless of the scale of their spatial resolution, however, it is important perceptually that these components' "common fate" over image motions be discriminable for figure-ground segregation. Despite the potential visual utility for maintaining synchronous information about the coherent motion of optical features at varying spatial scales, research suggests that features with different spatial frequency content are detected with different latencies, different temporal resolutions, and different contrast

sensitivities (see Kelly, 1985; Mather, 1987; Nakayama & Silverman, 1985).

A primary purpose of the present research was to evaluate the perception of coherent motion of spatially separated contours at multiple spatial scales. In our initial experiments, we tested the detection of motion of one or two separate contours that had the same spatial frequency content. We then compared this detection performance with performance in discriminating the correlated versus uncorrelated movements of a pair of contours whose spatial frequency contents either were the same or differed by more than two octaves. The results indicate that under a variety of viewing conditions, observers are surprisingly accurate at discriminating coherent motion of spatially separate contours, even when those contours differ in spatial frequency content.

GENERAL METHODS

Observers

Four adult observers participated in two or more of the three experiments. Two of the observers were authors (L.M. and R.B.), and the other 2 (C.R. and F.N.) were naive to the purposes of the experiments. All observers had normal visual acuity, and all except C.R. were experienced psychophysical observers. Each observer was given practice on all of the stimulus conditions prior to formal data collection.

Apparatus

The stimulus displays consisted of vertically oriented bars whose horizontal luminance profiles were defined by the difference of spatial Gaussian functions (DOG):

$$\text{DOG}(x) = 3\exp(-x^2/\sigma^2) - 2\exp(-x^2/2.25\sigma^2) + K,$$

where σ is the space constant in degrees visual angle adjusted for viewing distance and K is the mean background luminance (Badcock & Schor, 1985). This stimulus appears as a bright, blurred region flanked by two dark, blurred regions. The luminance profile of each DOG was generated by depositing the difference of Gauss-

This research was supported in part by NIMH Postdoctoral Training Grant MH15792 for L.M., NIH Grant EY07760 to R.B., NIH Grant EY05926 to J.S.L., and by NIH Vision Research Core Grant P30EY08126 to Vanderbilt University. The authors would like to thank J. Farley Norman and Cheryl Rye for participating in these experiments. Correspondence may be addressed to Lyn Mowafy, NASA-Ames Research Center, Mail Stop 262-3, Moffett Field, CA 94035. (E-mail: mowafy@eos.arc.nasa.gov)

ians pattern for a selected σ on a 126 horizontal $\times$ 60 vertical $\times$ 8 bit pixel region of the screen's bitmap. The maximum luminance value in the associated 256 gray-level color table was indexed to the center of the DOG's region of the bitmap. The remaining intensity values in this region then were adjusted to produce a DOG of the appropriate contrast. These new intensity values were written to a hardware lookup table that translated graphics memory contents into the intensity signals for video output. The stimuli were presented on an Apple monochrome monitor (640 horizontal $\times$ 480 vertical pixel resolution; 66.7 Hz noninterlaced frame rate) with a P4 phosphor. The average background luminance of the display was set to a midrange gray-level value and measured 10.1 cd/m². The displays were generated using a Macintosh II computer, which also controlled presentation of the trials and recorded the observers' responses. In the first two experiments, the observers viewed the displays binocularly in a darkened room from a distance of 430 cm; a small cross (3' arc) located in the center of the screen (2.8° $\times$ 2.13° arc) provided a fixation mark. Head position was constrained by a chinrest. The midpoint of each DOG was aligned vertically with the central fixation cross and was positioned horizontally 45' arc to the right or left of fixation for an on-center separation of 1.5° arc between bars at the start of each trial. At this viewing distance, the height of a bar was 16' arc.

Two spatial frequencies (2 and 12 cycles per degree [cpd]) were used in these experiments (the spatial extent of the negative Gaussian is 1.5 times wider than that of the positive Gaussian, and the total spatial extent of a DOG is the reciprocal of its center spatial frequency). The luminance contrast $(L_{\max} - L_{\min}) / (L_{\max} + L_{\min})$ of the 12-cpd DOGs was held constant at 50%. In the 2-cpd conditions, the contrast was either 50% or 5%. In these displays, contrast was defined mathematically on the basis of the gray-level value of the background. The intensity values for each DOG were adjusted relative to this midrange value. Any harmonic distortion in the DOGs due to z-axis nonlinearities in the luminances produced by the monitor would be minimal at these contrasts.

Apparent motion of a bar was created by changing the luminance values in the lookup table associated with the difference of Gaussians pattern to create a spatial phase shift; this occurred during the frame refresh flyback following every fourth refresh (60 msec). Perceptually, the bar appeared to change its spatial position smoothly. At the beginning of a movement sequence, the bar's initial position was always 45' arc from the fixation point. Thereafter, the direction and size of the displacement were selected at random from a uniform distribution of positions centered on the bar's initial position. The width of this distribution was determined by the bar's range of motion, which consisted of a set of possible positions the bar could assume during a given trial. In these experiments, the range of motion was varied between 15" and 75" arc, in 15" arc increments. Since successive positions were selected from the same range of motion, the probability distribution of the distance and direction of any given displacement depended on the position of the bar at that instant. These randomly changing positions were perceived as a sequence of horizontal displacements, randomly varying in both direction and distance.

Coherent (i.e., correlated) motion of a pair of bars was produced by selecting the same sequence of successive displacements for the two bars, so that both bars moved the same distance and in the same direction. Their relative spatial separation thereby remained constant. Uncorrelated motion was created by selecting each bar's successive positions independently. The bars could move in the same or opposite directions, and the size of their displacements was not necessarily equal (Figure 1). Thus, their relative spatial separation varied with each successive displacement.

Procedure

In each experiment, the observer fixated the cross in the center of the screen and triggered a trial with a keypress. A trial consisted

of two brief (240 msec) exposure intervals during which the DOGs displays were presented. In all experiments, the observer's task was a two-alternative forced-choice selection of the temporal interval containing the specified stimulus, guessing if necessary. Feedback was given after each trial.

EXPERIMENT 1

Movement Detection

The purpose of Experiment 1 was to determine observers' sensitivities to displacements of the bar(s). Although absolute sensitivity was not our primary interest in these studies, these measurements provided preliminary calibration of movement detectability for later comparison with the discrimination of correlated versus uncorrelated movements.

Method

For each of two spatial frequencies (2 and 12 cpd) at 50% contrast, psychometric functions for movement detection were obtained under three conditions: a single bar, a pair of bars whose movements were correlated, and a pair of bars with uncorrelated movements. In the single-bar condition, the stimulus appeared to the right or left of the fixation point in separate blocks of trials; the observer knew in advance where the bar would appear during a block of trials. Five ranges of motion (15", 30", 45", 60", and 75" arc) were tested in each condition. The stimuli were presented for two 240-msec intervals, during only one of which the bar(s) moved. The observer's task was to indicate which of the two intervals contained movement of the bar(s).

Each of the six experimental conditions (2 spatial frequencies $\times$ 3 conditions of movement) was presented over four blocks of 125 trials (25 trials at each range of motion, randomly ordered within the block). The order of the 24 trial blocks was randomly determined for each observer and distributed over a period of several days. For each of the six conditions, each observer provided 100 judgments at each range of motion. After completing these trials, the contrast of the 2-cpd bars was reduced to 5% for a final set of measurements in the double-bar correlated-motion condition.

Results and Discussion

Figure 2 shows that the observers were able reliably to detect motion displacements as small as 30" arc. Notice that, although there was no clear difference between the correlated and uncorrelated double-bar conditions, performance under both double-bar conditions tended to be somewhat better than performance under the single-bar condition. Such superiority would be expected on the basis of probability summation, since the observers had two opportunities to detect motion in the double-bar conditions. To assess the contribution of probability summation, we used Luce's (1963) choice theory to predict the difference in detectability of single- and double-bar movements, assuming that in the double-bar conditions the two movements were independently visible. (The details of the model are developed in the Appendix. We chose this approach because we wished to compare detection performance in Experiment 1 with the discrimination of correlated versus uncorrelated movements in Experiment 2, and few alternative models permit this comparison. In the present application, the choice theory predictions are es-

Figure 1. Frame sequence for correlated and uncorrelated motion. In correlated motion, the direction and magnitude of both bars' displacements are equal.

essentially the same as those derived from signal detection theory.)

Assuming that the distance measure of discriminability, $-\ln \eta$, is a linear function of the range of motion, this model may be used to describe the psychometric function relating detection accuracy to the range of motion. Thus, we used a chi-square minimization technique to estimate a single pair of slope and intercept parameters to describe the psychometric functions in all three detection conditions.¹ These two parameters were used to describe detection accuracy in 15 conditions for each observer. For purposes of graphic illustration, the same model was applied to the combined data from all 4 observers, and the results are shown in Figure 2.

The open squares in Figure 2 indicate the percentage of correct responses for the single- and double-bar conditions predicted from choice theory. As may be seen, the model predicted more gain due to probability summation than was actually observed. This is evident primarily in the single-bar conditions, where the predicted performance usually was lower than the obtained performance. The chi-square measures of deviation for each of the 4 individual observers in the 2- and 12-cpd conditions were statistically significant ($p < .05$) for all but one of the eight cases. The total chi-square value summed over the

4 observers was $\chi^2(52) = 141.33$ ($p < .01$) in the 2-cpd conditions and $\chi^2(52) = 93.98$ ($p < .01$) in the 12-cpd conditions. The fact that the increase in detectability of the double-bar patterns, relative to the single-bar patterns, was less than predicted indicates that the two concurrent movements were not independently detectable. That is, the two concurrent movements in the double-bar conditions apparently were perceptually partially correlated. This is consistent with the fact that performance in the correlated double-bar conditions usually was slightly lower than that in the uncorrelated conditions—as if the correlated movements were partially correlated perceptually, as well as physically. In any case, the obtained improvements in performance in the double-bar conditions apparently were due primarily to the effects of probability summation of visual information from partially independent sources.

Because the observers were more accurate in the 2-cpd conditions, they were retested on the correlated motion displays at 5% contrast. At the lower contrast, motion-detection performance was comparable to that measured with the high-contrast 12-cpd bars in the correlated-motion condition (Figure 3). This contrast difference was maintained in the subsequent experiments in order to measure discriminability of correlated-motion targets under con-

Figure 2. Results of Experiment 1 indicating accuracy for discriminating the interval containing movement of the bar(s) from a stationary interval.

ditions in which the bars' movements were approximately equal in their detectability.

EXPERIMENT 2

Discrimination of Correlated and Uncorrelated Movements

Having established the absolute sensitivity for detecting movements of bars viewed in the near periphery, we next sought to determine how accurately observers could discriminate correlated from uncorrelated motion of these bars at near-threshold displacements. Accurate performance on this discrimination would require, of course, that observers detect the motion of both bars and also determine whether the bars move in the same directions by the same amounts.

Method

DOGs at two spatial frequencies (2 cpd at 5% contrast and 12 cpd at 50% contrast) were paired to create three conditions: both 2 cpd, both 12 cpd, and a 2-12-cpd combination. In the combined spatial frequency condition, the two different spatial frequencies appeared on either side of the fixation cross equally often, and the observers

knew their relative locations at the beginning of a trial block. Discrimination accuracy was measured over five ranges of motion (15", 30", 45", 60", and 75" arc). In one temporal interval, the bars' movements were correlated; in the other interval, they were uncorrelated. The observer's task was to indicate in which interval the movements were correlated. Each of the three conditions was presented in four separate blocks of 125 trials (25 trials at each range of motion, randomly ordered). The 12 blocks of trials were randomly ordered for each observer and provided a total of 100 judgments by each observer at each range of motion in all conditions.

Results and Discussion

Figure 3 compares the detection of a pair of bars whose movements were correlated (from Experiment 1) with the discrimination between correlated and uncorrelated movements. For both spatial frequency conditions, performance on the detection task was consistently higher than was performance on the discrimination task. This difference would be expected, however, considering that detection of motion can be accomplished whenever the motion of either bar is seen. (Recall that the detection of the double-bar motion exceeded that of the single-bar motion by an amount similar to that expected from probability summa-

tion.) On the other hand, correct discrimination requires simultaneous detection of the movements of both bars.

To determine whether the accuracy in discriminating correlated versus uncorrelated movements was limited simply by the simultaneous detectability of the both contours' movements, we again used Luce's (1963) choice theory to evaluate and compare performance in these two tasks (see the Appendix). Using a chi-square minimization technique, we estimated a single pair of slope and intercept parameters to describe the psychometric functions relating range of motion to the detection of correlated double-bar movements (Experiment 1) and the discrimination of correlated versus uncorrelated movements (Experiment 2). These parameters related only to the detectability of movement, with no assumed limitation at all on the discriminability of the correlated versus uncorrelated movements. The predicted accuracies for the combined data of the 4 observers are indicated by the open squares (detection) and open circles (discrimination) in Figure 3.

Figure 3. Combined results of Experiments 1 and 2 indicating accuracy for detecting the movement of a pair of correlated bars and discriminating correlated movements of the bars from uncorrelated movements.

As may be seen, the model predicted accuracies that correspond quite well with both the obtained detection-accuracy data and the obtained discrimination-accuracy data, particularly in the 2-cpd condition. The chi-square measures of fit were not significant ($p > .10$) for any of the 4 observers in the 2-cpd condition, and the total chi-square value summed over all 4 observers was close to the expected value under the null hypothesis [$\chi^2(32) = 33.56$]. Small, but noticeable, discrepancies occurred between the predicted and obtained performance in the 12-cpd conditions. These discrepancies are attributable to 2 observers whose performance was less accurate and less consistent than that of the other 2 observers. For these 2 observers, the chi-square values were significant [$\chi^2(8) = 15.95$, $p < .05$, and $\chi^2(8) = 25.37$, $p < .01$]. However, for the other two observers, obtained performance in all 10 conditions was closely described by the model [$\chi^2(8) = 6.47$, $p > .50$, and $\chi^2(8) = 7.30$, $p > .50$]. Not surprisingly, the total chi square was statistically significant [$\chi^2(32) = 55.09$, $p < .01$]. Nevertheless, as is evident in Figure 3, the relative performance in the detection and discrimination tasks is consistent with the idea that the discriminability of correlated versus uncorrelated movements is limited only by the detectability of movement by both component contours.

Within the context of choice theory, any slight discrepancies between obtained and predicted discrimination performance may reasonably be attributed to either or both of two factors. First, this model employed a simplifying assumption to represent the movement of each of four bars (two in the correlated-motion interval and two in the uncorrelated-motion interval) as a single displacement to either the right or left. The movement of each bar on each trial, however, consisted of three successive displacements with independent random distances and directions. Thus, the uncorrelated motions often consisted of two displacements in the same direction and, sometimes, for the same distance. The representation of the uncorrelated movements as being composed of two equal and opposite displacements surely overestimates the perceptual discriminability of these stimuli.

Second, we assumed that the detectabilities of movements of the four bars on each trial were stochastically independent. The two component movements in each pattern were probably perceptually correlated to a slight degree and, therefore, less easily detected and identified than would be expected from the independence assumption. Evidence supporting this conjecture was obtained in the results of Experiment 1, where differences in performance between the single- and double-bar conditions were less than that predicted from independence and where the two correlated motions often were less accurately detected than were the two uncorrelated motions. With these considerations in mind, the similarities of the obtained and predicted discrimination accuracies suggest that errors in discriminating correlated from uncorrelated movements of these two bars derived mainly from limitations on the visibility of any movement at all rather than

from any appreciable limitations on the visibility of coherence as such.²

Figure 4 plots the psychometric functions for discriminating correlated movements from uncorrelated movements when the spatial frequencies of the DOGs were the same and when they were different. The overlap of the three functions indicates that the observers were able to discriminate correlated movements of the pair with dissimilar spatial frequencies and contrasts in the 2-12-cpd display about as accurately as they performed on the displays containing bars of the same spatial frequency and contrast.

Nevertheless, the three functions do appear to diverge at the longer ranges of motion. An analysis of variance indicated no significant main effect of spatial frequency condition, but there was a significant interaction of spatial frequency and range of motion [$F(8,24) = 2.32$, $p < .05$]. This divergence of the three functions, however, was not consistent among the 4 observers. In Figure 5, discrimination performance is plotted for each of the 4 observers. For 2 of the 4 observers (C.R. and F.N.), the three functions did not diverge as the range of motion increased. For the other 2 observers (L.M. and R.B.), the manner in which the psychometric functions diverged was not systematic. Insofar as performance in this discrimination task was limited by the visibility of the movements of both bars, performance in the 2-12-cpd condition would be expected to be no better than the minimum of that for either the 2- or 12-cpd conditions. This limitation is evident in the data for all 4 observers. Therefore, these results indicate that even when the two bars differed in spatial frequency by more than two octaves and in contrast by a factor of 10, discrimination accuracy remained comparable to that obtained when the contours were equal in spatial frequency and in contrast.

Figure 4. Results of Experiment 2 indicating accuracy for discriminating the correlated-motion interval from the uncorrelated interval for three spatial frequency combinations.

EXPERIMENT 3

Monocular Versus Dichoptic Viewing

Results from the first two experiments demonstrate that humans are remarkably adept at detecting coherent motion: If observers can detect that two spatially separated contours are moving, they can also tell whether the relative movements of those contours are correlated. This sensitivity to coherent motion might be affected, however, by the angular separation of the two contours or by presentation of the two contours separately to the two eyes. Angular separation is a relevant factor for two reasons. First, there are well-established changes in motion sensitivities with retinal eccentricity that could affect sensitivity to coherent motion (Harwerth & Levi, 1986; Tyler & Torres, 1972). Second, increasing lateral separation requires that local measurements be integrated over a greater angular extent of the visual field to yield the perception of globally coherent motion. Previously, we had found that correlated versus uncorrelated movements of three high-contrast points were highly discriminable (equal to the detection of any motion) even when the points were separated by up to 16° arc (Lappin, Wason, & Akutsu, 1987). We wondered whether a similar result might be obtained for a pair of low-contrast blurred contours.

Dichoptic presentation explores yet another conceivable determinant of coherent-motion perception. The detectability of coherent motion might diminish under dichoptic viewing conditions because small binocularly uncorrelated microsaccades might perturb the fidelity of the motion signals supporting coherence detection.

Method

For Experiment 3, we created a new stimulus display that presented 2-cpd DOGs (5% contrast) at a viewing distance of 212 cm. The vertical height of each bar was 32' arc. In the monocular viewing condition, the observer's head position was constrained using a chinrest. The observer viewed the monocular display with the preferred (right) eye. In the dichoptic viewing condition, the stimulus displays were constructed to present a fixation cross and one bar on each of two separate CRT screens. These displays were viewed through a mirror stereoscope with the head constrained. Fixation crosses in the dichoptic displays served to control eye position and to aid the observer in fusing the two displays.

In the monocular and fused dichoptic conditions, the spatial separation of the pair of bars was varied to correspond to 1°, 2°, 3°, and 4° separation. The two viewing conditions (monocular and dichoptic) and the four angular separations were combined factorially to create eight stimulus conditions. Two additional stimulus conditions were created at the 4° arc separation. In these conditions, the contrast of the bars was raised from 5% to 50%. This control was added to determine if any decline in motion sensitivity at the larger angular separations could be attributed to lower contrast sensitivity in the periphery. For each of the 10 conditions, discrimination thresholds were measured using two ranges of motion (1.5' and 3.0' arc).

The various conditions again were presented over several randomly ordered trial blocks. Observers L.M. and F.N. provided 150 discrimination judgments at each range of motion for both viewing conditions.

Figure 5. Individual observers' accuracy at discriminating the correlated-motion interval for three spatial frequency combinations.

Results and Discussion

Not surprisingly, discrimination performance declined with increasing angular separation of the two bars (Figure 6). This finding is consistent with the known fall-off in motion sensitivity (Leibowitz, Johnson, & Isabelle, 1972; Tyler & Torres, 1972) and in contrast sensitivity (Robson & Graham, 1981) with retinal eccentricity. Notice, however, that testing with DOGs of 50% contrast at 4° arc separation improved performance significantly. This improvement suggests that the decline in discrimination performance with increasing separation can be attributed at least in part to a decline in contrast sensitivity. The decline in discrimination accuracy also could be related to a spatial scaling of the stimulus, with retinal eccentricity and a corresponding decline in motion sensitivity to high spatial frequencies in the periphery. Regardless of the effectiveness of a specific spatial frequency in the visual field, however, the results indicate that, up to at least 4° arc, spatial separation of a pair of contours

appears to have no effect on the discriminability of coherent motion.

Of further interest is the fact that performance with dichoptic stimulation differed only slightly from that measured under monocular stimulation. In general, this pattern of results provides no support for the idea that microsaccadic eye movements, which are uncorrelated between the two eyes, introduce additional variability into the motion signals produced by these contours.

GENERAL DISCUSSION

The principal result of these experiments was that the observers could determine whether the movements of spatially separate contours were correlated even when those contours were imaged at different scales of spatial resolution. The observers were sensitive, in other words, to the constancy of the spatial distance between the moving contours. Of particular interest is the fact that they were

Figure 6. Results of Experiment 3 indicating accuracy for discriminating correlated motion as a function of angular separation under monocular and dichoptic viewing conditions.

able to judge coherent movements with about the same level of accuracy as they could detect any movement at all. (Our present application of Luce's choice theory indicates that discrimination performance was predictable entirely from the obtained detectability of movement in the two-bar detection conditions, with no additional limitation on the discriminability of the correlated versus uncorrelated movements.) This sensitivity to the coherent motion of contours at differing spatial scales is interesting considering that the putative neural mechanisms responsive to differing spatial frequencies also differ in their temporal resolution and response latencies. Most environmental objects, however, produce retinal images containing information at multiple spatial scales, so it is adaptive that the visual system can successfully integrate information at these multiple scales in a way that preserves the structural stability of moving objects.

Movements of the two spatially separate contours were temporally synchronized in the present experiments. The temporal synchrony of these movements may play an important role in coordinating and integrating visual information about the relative spatial structure and movements

of contours at separate locations and differing spatial scales. Indeed, the problem of how to obtain coherent structural information from optical image-processing mechanisms at multiple scales of temporal (and spatial) resolution constitutes a fundamental and difficult theoretical problem in image processing. Koenderink (1988) has pointed out that a solution to this theoretical problem can be obtained by designing mechanisms with differing time constants so as to synchronize information about the present moment while integrating over different durations of the immediate past. Neurophysiological evidence that the vertebrate visual system may be designed this way has been obtained by Stanford (1987), who showed that responses in ganglion cells originating in varying regions of the cat's retina are temporally synchronized despite potential differences in spatial and temporal resolution. The results of the present experiments are compatible with the hypothesis that human vision maintains temporal synchrony of information about moving contours at multiple spatial scales, although this hypothesis was not tested directly by varying the temporal phase relations among the separate moving contours. We are now completing a study of this question that will be reported in a forthcoming paper.

Although we are only beginning to understand synchronicity among channels processing concurrent-motion information, our results may have interesting implications for models of motion perception. For example, a recent computational model (Yuille & Grzywacz, 1988) proposes that information about coherent visual motion is extracted in two stages of processing. An initial stage performs a velocity measurement within local regions of the visual field; the second, so-called "smoothing," stage uses these local measurements to construct a velocity field over an entire display. Our results suggesting that coherent motion is discriminable when motion is detected would imply that either this second stage operates with extremely high fidelity, utilizing essentially all the information available from local measurements, or that relative motion information may not be processed in separate stages. Of course, our two-bar displays are quite simple compared with the complex motion fields encountered in nature, so one might argue that our results do not tax the computational coherence theory. However, other work in our lab (Lappin, Norman, & Mowafy, 1989, in press) indicates that coherence in multielement displays is detectable whenever motion is seen and that these coherent motions include geometric transformations other than the horizontal translations examined in the present experiments.

And finally, we were surprised by the finding in Experiment 3 that coherent motion was easily detectable even when the two moving contours were displayed dichoptically. Although these are preliminary results from a limited range of conditions, they suggest interesting questions about the temporal coordination between spatio-temporal signals from the two eyes. Considered together with our finding that the ability to detect coherence may be essentially equivalent to the ability to detect movement,

it would suggest that processing relative movements in the environment is a fundamental characteristic of human motion perception.

REFERENCES

- BADCOCK, D. R., & SCHOR, C. M. (1985). Depth-increment detection function for individual spatial channels. *Journal of the Optical Society of America*, **A2**, 1211-1215.
- HARWERTH, R. S., & LEVI, D. M. (1986). Reaction time as a measure of suprathreshold grating detection. *Vision Research*, **18**, 1579-1586.
- KELLY, D. H. (1985). Visual processing of moving stimuli. *Journal of the Optical Society of America*, **A2**, 216-225.
- KOENDERINK, J. J. (1988). Scale-time. *Biological Cybernetics*, **58**, 159-162.
- LAPPIN, J. S., NORMAN, J. F., & MOWAFY, L. (1989). Detecting geometric components of optic flow. *Investigative Ophthalmology & Visual Science* (Suppl. 30), 264.
- LAPPIN, J. S., NORMAN, J. F., & MOWAFY, L. (in press). The detectability of geometric structure in rapidly changing optical patterns. *Perception*.
- LAPPIN, J. S., WASON, T. D., & AKUTSU, H. (1987). Visual detection of common motion of spatially separate points. *Bulletin of the Psychonomic Society*, **25**, 342.
- LEGGE, G. E., & CAMPBELL, F. W. (1981). Displacement detection in human vision. *Vision Research*, **21**, 205-213.
- LEIBOWITZ, H. W., JOHNSON, C. A., & ISABELLE, E. (1972). Peripheral motion detection and refractive error. *Science*, **177**, 1207-1208.
- LUCE, R. D. (1963). Detection and recognition. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (Vol. 1). New York: Wiley.
- MATHER, D. (1987). The dependence of edge displacement thresholds on edge blur, contrast, and displacement distance. *Vision Research*, **27**, 1631-1637.
- NAKAYAMA, K., & SILVERMAN, G. H. (1985). Detection and discrimination of sinusoidal grating displacements. *Journal of the Optical Society of America*, **A2**, 267-274.
- ROBSON, J. G., & GRAHAM, N. (1981). Probability summation and regional variation in contrast sensitivity across the visual field. *Vision Science*, **21**, 409-418.
- STANFORD, L. R. (1987). Conduction velocity variations minimize conduction time differences among retinal ganglion cell axons. *Science*, **238**, 358-360.
- TYLER, C., & TORRES, J. (1972). Frequency response characteristics for sinusoidal movement in the fovea and periphery. *Perception & Psychophysics*, **12**, 232-236.
- YUILLE, A. L., & GRZYWACZ, N. M. (1988). A computational theory for the perception of coherent motion vision. *Nature*, **333**, 71-74.

NOTES

1. We are grateful to David Gilden for his help in designing the computer program to perform this minimization procedure.
2. In other experiments involving longer sequences of random two-dimensional translations (Lappin et al., 1987) and rotations (Lappin et al., 1989) of high-contrast points, we have found that the discriminability of correlated versus uncorrelated movements has equaled the detectability of movement. These results indicate that the concurrent movements of multiple points are correlated visually. Precisely why the movements of the high-contrast points are correlated visually while the movement of low-contrast DOGs are more nearly independent is not yet known.

APPENDIX

The purpose of this Appendix is to describe rationale and a quantitative model for evaluating the performance in discriminating correlated motions from uncorrelated motions of two bars by comparison with the performance in detecting the motion of either bar. The principal assumption adopted as a basis for this

evaluation was that errors of discrimination resulted entirely from errors in detecting the motion of either one or both bars. This assumption seemed compatible with the subjective difficulties in discriminating the small displacements of these low-contrast bars. Recall that performance in detecting motion in the double-bar conditions was better than that in the single-bar conditions by an amount comparable to what would have been expected from the summation of partially independent visual information about the motion of each of the two bars. Thus, if visual information about these two motions was indeed independent, then the accuracy in discriminating whether the two motions are the same or different should be lower than that in detecting the motion of either one or both bars, because accurate discrimination obviously would require detection of both bars, whereas detection performance in the double-bar conditions would require only the detection of either one, or the other, or both bars.

As an initial simplifying assumption, we considered this discrimination task as a *same/different* task, involving the comparison between two motions composed of a single displacement to either the right (*r*) or the left (*l*). That is, we supposed that each stimulus presentation was one of four equally likely alternative events: *rr*, *ll*, *rl*, and *lr*, where the first two alternatives are *same* and the latter two *different*. The psychophysical procedure in this experiment was a two-alternative forced-choice task, involving two such stimuli, one of which was *same* and the other *different*. Thus, the stimulus event on each trial was one of eight equally likely alternatives: (*rr*, *rl*), (*rr*, *lr*), (*ll*, *rl*), (*ll*, *lr*), (*rl*, *rr*), and so on. We employed Luce's (1963) choice theory to represent the conditional probability that a given stimulus, s_i , will lead to the response, R_j , associated with a particular stimulus, s_j . The fundamental assumption of choice theory is that this conditional probability equals the ratio of the scale value for the perceptual similarity of this pair of stimuli relative to the sum of the scale values for all the potential stimulus pairs that include s_i . That is, in a complete identification experiment if response R_j is assigned to stimulus s_j , then the conditional probability of response R_j , given stimulus s_i , is represented as

$$P(R_j|s_i) = \eta_{(i,j)} / \sum_k \eta_{(i,k)}, \quad (1)$$

where $\eta_{(i,j)}$ is the similarity scale value for the pair of stimuli (s_i, s_j).

Now consider a detection experiment involving three alternative stimuli, consisting of motion to the right, to the left, or no motion (*r*, *l*, and *n*, respectively). We assume that the stimuli *r* and *l* constitute independent properties, perceptually distinguished from the stationary stimulus *n* by orthogonal vectors of equal length. That is, if $d(r, l)$ is the perceptual distance between the stimulus pair *r* and *l*, and if $d(r, n)$ and $d(l, n)$ are the corresponding distances of *r* and *l* from the stationary stimulus *n*, then we assume that

$$d(r, l) = \sqrt{[d(r, n)^2 + d(l, n)^2]} = \sqrt{2}d(r, n) = \sqrt{2}d(l, n) \quad (2)$$

If the similarity scale value for the pairs (*r*, *n*) and (*l*, *n*) is η , and if the scale values for the pairs (*r*, *r*), (*l*, *l*), and (*n*, *n*) are set at 1, then the scale value for the pair (*r*, *l*) equals η^2 . The latter value is associated with the choice theory measure of distance as $d(s_i, s_j) = -\log \eta(s_i, s_j)$. See Luce (1963) for a more complete exposition.

Thus, for the single-bar forced-choice detection task in Experiment 1, we can construct the scale values for the set of stimulus-response pairs as follows:

Stimulus Presentation	Response			
	1		2	
	(r,n)	(l,n)	(n,r)	(n,l)
(r,n)	1	η^2	η^2	η^2
(l,n)	η^2	1	η^2	η^2
(n,r)	η^2	η^2	1	η^2
(n,l)	η^2	η^2	η^2	1

(3)

Accordingly, by combining Equations 1 and 3, the probability of a correct response in this single-bar two-alternative forced-choice detection task is given as

$$P_1 = (1 + \eta^2)/(1 + 3\eta^2), \quad (4)$$

solving for the parameter η , we have

$$\eta = [(1 - P_1)/(3P_1 - 1)]^{1/2}. \quad (5)$$

Now we follow similar procedures to construct the similarity scale values for the double-bar detection task of Experiment 1. For the conditions in which the two bars were correlated, we again have four alternative stimulus presentations (rr, nn), (ll, nn), (nn, rr), and (nn, ll). (For the uncorrelated double-bar conditions, the construction is essentially the same, of course, except that the two moving stimuli are $[rl]$ and $[lr]$.) Accordingly, the similarity scale values for the first row of the stimulus-response matrix in this double-bar detection task is given as

Stimulus Presentation	Response			
	1		2	
	(rr,nn)	(ll,nn)	(nn,rr)	(nn,ll)
(rr,nn)	1	η^2	η^2	η^2

(6)

with the remaining rows containing permutations of the same values. Following the same procedure as that for the single-bar detection task we find that

$$P_2 = (1 + \eta^2)/(1 + 3\eta^2) \quad (7)$$

and

$$\eta = \sqrt{[(1 - P_2)/(3P_2 - 1)]}. \quad (8)$$

Finally, we construct the similarity scale values for the first row of the stimulus-response matrix in the discrimination experiment:

Stimulus Presentation	Response							
	1				2			
	(rr,rl)	(rr,lr)	(ll,rl)	(ll,lr)	(rl,rr)	(lr,rr)	(rl,ll)	(lr,ll)
(rr,rl)	1	η^2	η^2	η^2	η^2	η^2	η^2	η^2

(9)

and again the remaining seven rows follow the same pattern as permutations of the same set of scale values. The probability of a correct response in this forced-choice *same/different* discrimination task is then given as

$$P_{SD} = (1 + 2\eta^2 + \eta^2\sqrt{2})/(1 + 6\eta^2 + \eta^2\sqrt{2}). \quad (10)$$

Thus, Equations 4, 7, and 10 represent the observed accuracies in the three tasks—single-bar detection, double-bar detection, and discrimination, respectively—as functions of the same underlying visual detectability parameter, η .

An additional constraint on the theoretical representation on the performance of these three tasks pertains to the effects of variations in the range of motion. If we assume that detectability of movement of any given contour is a linearly increasing function of the range of motion, then the relative performance for the five different ranges used in these experiments may be described by the two parameters of this linear function. Specifically, we can write

$$-\log \eta = \gamma R + \theta, \quad (11)$$

where R is the range of movement, γ is the gain of the movement detection system, and θ is an additive constant. Thus, for example, the two parameters γ and θ , along with Equations 4 and 7, were used to predict the relative detection performance in the 15 independent detection conditions of Experiment 1 (five ranges of motion each for the single-bar and two double-bar conditions).

These parameter estimates were obtained by an iterative chi-square minimization routine. The chi-square measure of deviation between the obtained and predicted percentages of correct responses was

$$\chi^2 = \sum_i [N(O_i - P_i)^2 / P_i(1 - P_i)], \quad (12)$$

where N is the number of trials, O and P are the obtained and predicted percentages of correct responses, respectively, and the denominator is based on the variance of the binomial distribution.

To summarize, the predictions of this model were based on the following assumptions: (1) Discrimination accuracy was limited entirely by the detectabilities of the motions of the four bars on each trial (two bars in each interval). That is, the single free parameter in the model measured the detection of motion versus no motion of a single bar. (2) The detectabilities of the four bars' movements on each trial were stochastically independent. (3) The motion of a given bar on a given trial was represented as a single displacement either to the right or to the left. Thus, correlated motions were represented as displacements of both bars in the same direction; uncorrelated motions were represented as displacements of the two bars in opposite directions, one right and one left. (4) Visual information about these two opposite motions was assumed to be stochastically independent. Thus, the discriminability of motion to the right and left was assumed to be better than the detectability of each from no motion by a factor of $\sqrt{2}$ —as if the two opposite movements were represented by two orthogonal vectors representing the discriminability of a given direction of motion from no motion.