Temporal factors in the discrimination of coherent motion

LYN MOWAFY, JOSEPH S. LAPPIN, BARTON L. ANDERSON, and DEBORAH L. MAUK
Vanderbilt University, Nashville, Tennessee

When observers view the relative movements of a pair of bars defined by the difference of spatial Gaussian functions (DOGs), they can accurately discriminate coherent movements over a range of temporal frequencies and temporal asynchronies. Of particular interest is the fact that performance accuracy is maintained even when the two bars differ in spatial-frequency content and contrast. 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 observation interval, both elements moved in the same direction and by the same magnitude (correlated), and in the other interval, the movements were independent (uncorrelated). Temporal asynchronies were introduced by delaying the displacement of the right bar relative to that of the left bar in each interval. Observers were able to discriminate correlated versus uncorrelated movements up to a 45-60-msec temporal delay between the two elements' relative displacements. If motion processing is accomplished by mechanisms operating over multiple spatial and temporal scales, the visual system's tolerance of temporal delays among correlated signals may facilitate their space-time integration, thereby capitalizing on the perceptual utility of coherent-motion information for image segmentation and interpolating surface structure from the movements of spatially separated features.

Human observers are remarkably sensitive to the coherence, or correlation, among concurrent displacements of moving elements (cf. Lappin, Norman, & Mowafy, 1991; Lappin, Wason, & Akutsu, 1987; Mowafy, 1991). Previously we reported the results of a series of experiments suggesting that observers could discriminate whether the movements of spatially separate contours were correlated even when those contours were imaged at different scales of spatial resolution and contrast. Moreover, they were able to judge coherent movements with about the same level of accuracy as they could detect any movement at all (Mowafy, Blake, & Lappin, 1990). Since it is generally agreed that local motion detection is accomplished by a set of independent mechanisms selectively tuned to different spatial-frequency bandwidths (Adelson & Bergen, 1985; Heeger, 1987; van Santen & Sperling, 1985; Watson & Ahumada, 1985; Wilson & Bergen, 1979), this sensitivity to the coherent movements of contours at differing spatial scales presents an interesting problem: How does the visual system integrate the local signals for the perception of globally coherent motion? Our results, as well as evidence of cooperativity in the motion-processing system (Williams, Phillips, & Sekuler, 1986; Williams & Sekuler, 1984) and of mo-

tion capture (Ramachandran & Anstis, 1983), would suggest that at an early processing stage the system successfully integrates information with exceptionally high fidelity over large areas of the retina and at multiple spatial scales. On the other hand, the well-known "aperture problem" demonstrations have shown that the perception of coherent global motion is disrupted when a "plaid" pattern is created by the superimposition of a pair of moving gratings differing in their spatial-frequency contents (Adelson & Movshon, 1982) and their contrast (Stone, Watson, & Mulligan, 1990). These apparently conflicting results suggest that although coherent-motion perception appears to involve fundamental visual processes, there also are conditions under which global coherence is not perceptible.

In the present experiments, we sought to explore further the conditions under which human observers are capable of accurately discriminating coherent motion of a pair of spatially separate contours when those contours are imaged at the same or different scales of spatial resolution. Specifically, we wondered whether the temporal synchrony of the movements would be a critical determinant in integrating visual information about the relative spatial structure and movements of these contours. Temporal synchrony of the movements is of interest because there is considerable evidence to suggest that the neural mechanisms responsive to differing spatial frequencies also differ in their temporal response properties (cf. Anderson & Burr, 1985; Georgeson, 1987; Kelly, 1979; Watson, Thompson, Murphy, & Nachmias, 1980). In a similar vein, computational analysis of the problem of deriving global motion from local velocity measures

This research was supported in part by NIMH Postdoctoral Training Grant MH15792 to L.M., NIH Grant EY05926 to J.S.L., and NIH Vision Research Core Grant P30EY08126 to Vanderbilt University. The authors would like to thank Sabrina Hitt for participating in these experiments. Correspondence may be addressed to L. Mowafy, University of Dayton Research Institute, P.O. Box 2020, Higley, AZ 85236-2020.

would suggest that, constrained by the computational complexity alone, the visual system should integrate information at multiple spatial scales only if the signals are synchronized temporally (cf. Watson & Ahumada, 1985; Wilson & Bergen, 1979).

Despite the apparent limitations imposed on the system for tolerating temporal asynchronies, the available psychophysical evidence is ambiguous as to the effect of temporal synchrony in coherent-motion perception. On the one hand, it has been demonstrated that human observers are very sensitive to temporal asynchronies. For example, McKee and Welch (1985) found that observers could discriminate velocities based on the detection of tiny (1-msec) temporal asynchronies in a motion sequence. Nakayama and Silverman (1984) also found that the upper displacement limit for motion-based image segmentation is doubled given a brief (10-msec) temporal delay between signals. These results indicate that the visual system is remarkably sensitive to temporal asynchrony and can use this information for velocity discrimination and image segmentation. Given the system's sensitivity to temporally asynchronous signals, it is feasible that temporal asynchrony among spatially separated contours could disrupt the observer's ability to perceive their spatially coherent movements. On the other hand, there also is evidence to suggest that the system may be relatively unselective in its temporal requirements when discriminating motion coherence of image features. Silverstein, Klein, and Carney (1991) have reported that although observers are sensitive to temporal asynchrony of as little as 1 msec in a vernier target undergoing linear motion (4 msec for rotary motion), phase offsets as large as 16 msec were required to discriminate changes in the spatial separation of two collinear targets. These results would suggest that temporal asynchrony among spatially separated contours may not disrupt the perception of their spatially coherent movements.

In the present experiments, we tested performance in discriminating the correlated versus uncorrelated movements of a pair of bars when the displacement of one of the bars was delayed relative to that of the other. The results indicate that temporal delays between the movements of these spatially separate elements do not completely disrupt coherent-motion perception. Observers can accurately discriminate coherent motion even with a relatively large temporal asynchrony between the incoming signals. Moreover, the spatial-frequency content of the elements has virtually no effect on the ability to discriminate globally coherent movements.

GENERAL METHOD

Observers

Four adult observers participated in these experiments; 3 of the observers were authors (L.M., B.A., and D.M.), but B.A., D.M., and the 4th observer (S.H.) were naive to the purposes of the experiments during data collection. All observers had normal visual acuity, and all except S.H. 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):

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

where s is the space constant in degrees of visual angle adjusted for viewing distance and K is the mean background luminance. The luminance profile of each DOG was generated by depositing the pattern of the difference of Gaussians for a selected s on a 126 h $\times$ 60 v $\times$ 8 bit pixel region of the screen's bitmap with the maximum luminance value indexed to the DOG's center. The remaining intensity values in this region then were adjusted to produce a DOG of the appropriate contrast (details on the construction of the stimulus displays have been described previously; see Mowafy et al., 1990). The stimuli were presented on an AppleColor high-resolution RGB monitor (640 h $\times$ 480 v pixel resolution; 66.7-Hz noninterlaced frame rate). Each DOG appeared as a bright, blurred region flanked by two dark, blurred regions. 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.

The observers viewed the displays binocularly in a darkened room from a distance of 430 cm (a front-surface mirror was used to achieve this viewing distance). Head position was constrained by a chinrest, and a small cross (3' of arc) located in the center of the screen (2.8° $\times$ 2.13° of arc) provided a fixation mark. The midpoint of each DOG was aligned vertically with the central fixation cross, and it was positioned horizontally 45' of arc to the right or left of fixation, for an on-center separation of 1.5° of arc between bars at the start of each trial. At this viewing distance, the height of a bar was 16' of arc.

DOGs at two spatial frequencies (2 cycles per degree [cpd] and 12 cpd) were paired to create three conditions: both 2 cpd, both 12 cpd, and a combination of 2 and 12 cpd. The luminance contrast $(L_{\max} - L_{\min}) / (L_{\max} + L_{\min})$ of the 12-cpd DOGs was held constant at 50%, and the 2-cpd DOGs were held constant at 5% contrast. This contrast difference was used in order to measure the discriminability of correlated motion targets under conditions in which the bars' movements were approximately equal in their detectability. (Our previous research had demonstrated that motion-detection performance was comparable in the two spatial-frequency conditions when the contrast of the 2-cpd DOGs was reduced by a factor of 10; Mowafy et al., 1990.) 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.

Apparent motion of a bar was produced by a spatial phase shift of its luminance distribution at a constant (experimentally manipulated) temporal frequency and occurred during the frame-refresh flyback. At the beginning of a movement sequence, the bar was positioned 45' of arc from the fixation point. Thereafter, the direction and size of its displacement were selected at random from a uniform distribution of positions centered about 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 or motion was varied between 15 and 75 sec of arc in increments of 15 sec of arc, yielding five range-of-motion conditions (15, 30, 45, 60, and 75 sec of arc). These shifts 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 suc-

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

cessive positions independently. The bars could move in the same or opposite directions, and the sizes of their displacements were not equal (Figure 1). Thus, their relative spatial separation varied with each successive displacement.

Procedure

At the beginning of each trial, the observer fixated the cross in the center of the screen and initiated the trial with a keypress. A trial consisted of two 1.0-sec exposure intervals during which the DOGs' displays were presented. In one temporal interval the bars' movements were correlated, while in the other interval they were uncorrelated. The observer's task was a two-alternative forced-choice selection of the temporal interval containing the correlated movements, and the observers were to guess, if necessary. Feedback was given after each trial.

EXPERIMENT 1

The Effect of a Temporal Phase Shift

In Experiment 1, we sought to determine how accurately observers could discriminate correlated from uncorrelated movements when the temporal phase relationships of those movements were shifted to create temporal asynchronies. A temporal phase shift in a sequence of coherent movements does not alter the correlation between the sizes and directions of the contours' successive spatial displacements. Rather, it introduces a temporal delay in their relative displacements so that one contour's spatial displacement is delayed relative to the correlated displacement of the other contour.

Method

In this experiment, the spatial phase shift producing the apparent motion of a bar (see General Method) occurred at a temporal frequency of 16.7 Hz. At 16.7 Hz, each bar underwent four 15-msec (total = 60 msec) vertical retraces at each spatial position during a given movement sequence. Temporal phase shifts were created by delaying the displacement of the right bar relative to that of the left bar, synchronizing this delay with the vertical-retrace flyback. This produced the five temporal-phase-shift conditions (0, $\frac{1}{4}$, $\frac{1}{2}$, $\frac{3}{4}$, and 1 cycle) shown in Figure 2. Each of the phase shifts was introduced to both the correlated and uncorrelated movements' intervals.

The five temporal-phase-shift conditions were combined factorially with the three spatial-frequency combinations (2–2 cpd, 12–12 cpd, and 2–12 cpd) to produce 15 experimental conditions. Each of the 15 experimental conditions was presented over 2 blocks of 125 trials (25 trials at each of the five ranges of motion, randomly ordered within a block). The order of the 30 trial blocks was randomly determined for each observer and distributed over a period of several days. Every observer provided 50 judgments for each range of motion in each of the 15 experimental conditions.

Results

Figure 3 compares the mean accuracy of the 4 observers at discriminating between correlated and uncorrelated movements for each of the 15 experimental conditions. The results for each of the three spatial-frequency conditions are shown separately for each of the temporal-phase-shift conditions. In the zero-phase-shift condition (top panel), the DOGs' movements were correlated spatially and temporally. It is evident from the overlap of the plots for the three spatial-frequency conditions that the observers were able to discriminate correlated movements of the pair of bars with dissimilar spatial frequencies (2–12 cpd) and contrasts as accurately as they performed on the displays containing bars of the same spatial frequency and contrast. Similar results were found in each of the remaining four temporal-phase-shift conditions (the spatial frequency of the delayed bar in the combined-spatial-frequency conditions also had no effect on performance accuracy). This replicates and extends our previous results: the relative spatial-frequency contents and contrasts of a pair of moving contours have virtually no effect on the visual system's capacity to discriminate coherent motion (Mowafy et al., 1990).

Not surprisingly, there was a gradual decline in performance as the temporal delay of the right bar's displacement increased with the temporal phase shift. Performance remained quite accurate, however, even with a $\frac{1}{2}$ -cycle temporal phase shift. That is, despite delays as long as 30 msec between the relative displacements of the right and left bars, the observers remained remarkably accurate at discriminating correlated versus uncorrelated movements. All 4 observers produced similar patterns of results. In Figure 4, mean performance over the five ranges of motion is presented as a function of the temporal phase shift for each observer. It is evident in this figure that despite some individual differences, discrimination accuracy was well above chance for all of the observers up

Figure 2. Five temporal-phase-shift conditions for a temporal frequency of 16.7 Hz. Each displacement of the right bar was delayed relative to the corresponding displacement of the left bar by an amount determined by the vertical-retrace period of the CRT. Note the actual duration of each temporal-phase-shift condition.

to a $\frac{1}{2}$ -cycle temporal phase shift in each of the spatial-frequency conditions. These results suggest that temporal delays between the relative movements of spatially separate contours do not completely disrupt coherent-motion perception, despite differences in their spatial-frequency contents and contrasts.

EXPERIMENT 2

The Effect of Temporal Frequency

Results of Experiment 1 demonstrated that humans are very sensitive to coherent movements, even when the moving elements differ in their spatial-frequency content, contrast, and temporal phase relationship. We wondered if this capacity to discriminate coherent motion would generalize to other temporal frequencies. Varying the temporal frequency of the contours' displacements explores two possible determinants of coherent-motion perception. First, since the separate channels tuned to different spatial frequencies also vary in their temporal resolution (An-

derson & Burr, 1985), temporal frequency may differentially affect performance. There is some evidence to suggest that sensitivity to sinusoidal oscillations does vary as a function of temporal frequency. Regan and Beverley (1978) compared motion sensitivity to in-phase and antiphase oscillations of the vertical edges of a square. For both motion types, sensitivity rose to a broad peak between 0.6 and 4 Hz. Silverstein et al. (1991) compared sensitivities to a small temporal phase shift in the oscillatory movements collinear with a pair of dots and oscillatory movements perpendicular to the pair of dots. They reported a peak sensitivity at about 3 Hz. Second, the actual duration of a temporal phase shift is dependent upon the temporal frequency of the movement sequence. At lower temporal frequencies, a $\frac{1}{2}$ -cycle phase shift would cause a longer temporal delay between the relative spatial displacements of the two contours. Conversely, a higher temporal frequency would decrease the delay at the same phase shift. The visual system's capacity to perceive coherent motion may be as much a function of the

Figure 3. Results of Experiment 1 ($N=4$) indicating accuracy for discriminating the interval containing the correlated movements of the pair of DOGs. Each plot shows discrimination accuracy for the three spatial-frequency conditions at a specific temporal phase shift.

Figure 4. Summary of results from Experiment 1 for individual observers. Each data point represents the mean over five ranges of motion at each temporal phase shift (250 observations).

actual duration of the delay as it is a function of the temporal phase relationship in the displacement sequence.

Method

Two new temporal-frequency conditions (8.3 and 33.4 Hz) were used in this experiment. At 8.3 Hz, each bar underwent eight 15-msec (total = 120 msec) vertical retraces at each spatial position. As shown in Figure 5, compared with the 16.7-Hz condition in Experiment 1, this effectively doubled the actual temporal delay of the right bar's displacement relative to the left bar for each temporal phase shift. For this temporal frequency, the five phase shifts (0, $\frac{1}{4}$, $\frac{1}{2}$, $\frac{3}{4}$, and 1 cycle) were combined factorially with the three spatial-frequency combinations (2-2 cpd, 12-12 cpd, and 2-12 cpd) to produce 15 experimental conditions. On the other hand, at 33.4 Hz, a bar maintained each position for only 30 msec (two 15-msec vertical retraces), so the actual temporal delay of the right bar was reduced to one half the delay for comparable phase shifts at 16.7 Hz. Since there were only two vertical-retrace periods at this temporal frequency, performance in the three spatial-frequency combinations could be measured only at temporal phase shifts of 0, $\frac{1}{2}$, and 1 cycle (Figure 6).

Each of the 15 experimental conditions at 8.3-Hz temporal frequency and 6 experimental conditions at 33.4 Hz was presented over two blocks of 125 trials (25 trials at each range of motion, randomly ordered within the block). The order of the trial blocks within each temporal-frequency condition was randomly determined for each observer and distributed over a period of several days. Observers B.A., D.M., and S.H. completed the 8.3-Hz condition before initiating the trials in the 33.4-Hz condition. Observer L.M. completed the trial blocks for the two temporal-frequency conditions in random order. This ordering of the trial blocks produced no differences in performance. Every observer provided 50 judgments for each of five ranges of motion in all conditions.

Results

Performance accuracies at each range of motion in the two new temporal-frequency conditions were very similar to those found in the 16.7-Hz condition in Experiment 1 (Figure 3). Therefore, for purposes of graphic illustration, mean performance over the five ranges of motion was computed and plotted as a function of tem-

Figure 5. Five temporal-phase-shift conditions for a temporal frequency of 8.3 Hz. Each displacement of the right bar was delayed relative to the corresponding displacement of the left bar by an amount determined by the vertical-retrace period of the CRT. Note that the actual duration of each temporal phase shift is twice that of the same phase shift at 16.7 Hz in Experiment 1.

poral phase shift for temporal frequencies of 16.7 Hz (Experiment 1) and 8.3 and 33.4 Hz (Experiment 2) in Figure 7. As we expected, again there were no systematic differences in performance in the three spatial-frequency conditions. The ability to perceive coherent movements of a pair of contours with the same or different spatial-frequency contents and contrasts was not affected by the temporal frequency of those movements.

As in Experiment 1, accuracy at discriminating the correlated movements was affected by the temporal phase shift. In the two new temporal-frequency conditions, again there is a decline in performance as the right bar's displacement is delayed $\frac{1}{4}$ cycle, $\frac{1}{2}$ cycle, and, eventually, 1 cycle relative to the left bar's displacement. Performance was not a direct function of temporal phase shift, however. For example, at $\frac{1}{2}$ -cycle temporal phase shift, performance is comparable in the 33.4- and 16.7-Hz conditions, but it is considerably lower in the 8.3-Hz condition. This would suggest that the visual system's sensitivity to coherent motion may be related more to the temporal de-

lay between the two contours' movements than to their temporal phase relationship.

As there were no differences in performance in the three spatial-frequency conditions, their means for each temporal frequency are plotted in Figure 8 as a function of the actual temporal delay. It is evident in this figure that up to a 45–60-msec temporal delay, the slopes of the three functions are approximately the same. They differ mainly in the magnitude of the number of correct responses. This magnitude difference reflects the fact that at any temporal delay, the phase shift is larger with the higher temporal frequencies. The slopes of the functions in the 16.7- and 8.3-Hz conditions decrease when the temporal delay exceeds 45 msec: the observers responded at chance performance.

This would suggest that there may be an upper limit to the temporal delay between the movements of spatially separate elements, beyond which their coherence is no longer perceptible. In Figure 9, performance for each of the individual observers in the three temporal-frequency

Figure 6. Three temporal-phase-shift conditions for a temporal frequency of 33.4 Hz. Note that the actual duration of each temporal phase shift is one half that of the same phase shift at 16.7 Hz in Experiment 1.

conditions is plotted as a function of the temporal delay between the two contours' movements. Although there were some individual differences, all 4 observers showed similar patterns of performance accuracy, including the slope change beyond the 45-msec delay.

At present, there is no reason to suspect that the 45-msec delay found in these experiments represents an absolute processing limit in the visual system. Of primary interest is the fact that performance remained above chance for relatively large temporal delays over a range of temporal frequencies. Thus, the visual system is tolerant of temporal asynchrony between the correlated spatial displacements of spatially separate contours, apparently disregarding differences in their spatial frequency and contrast as well as their displacement rate.

GENERAL DISCUSSION

The principal findings of these experiments are (1) that the capacity to perceive coherent movements between a pair of spatially separate contours is maintained over a

range of temporal frequencies, regardless of their spatial-frequency content and contrast, and (2) that the visual system can tolerate temporal delays of several milliseconds between the movements before their coherence is no longer perceptible.

Our finding that temporal frequency does not seem to affect coherent-motion perception would appear to conflict with the results of Regan and Beverley (1978) and Silverstein et al. (1991), who found that sensitivities to coherent oscillations of bars varied as a function of the temporal frequency. These authors have reported maximal sensitivity to coherent movements in the range of 1–4 Hz. The range of temporal frequencies tested here was 8–33 Hz, several orders of magnitude higher than those used previously. Although this could be a factor, we suspect that the discrepancy between these apparently conflicting results is probably related to the type of motion used in the stimulus displays. In the present studies, the DOGs' movements consisted of a series of discrete displacements in discrete time frames. Although the displacements occurred at a specific temporal rate, their dis-

Figure 7. Combined results of Experiments 1 and 2 indicating accuracy at discriminating the correlated motion interval for the three spatial-frequency conditions with temporal phase shifts at 16.7 Hz (Experiment 1) and 8.3 and 33.4 Hz (Experiment 2). Each data point represents the mean for 4 observers over five ranges of motion (1,000 observations).

crete nature introduced a broad spectrum of sinusoidal temporal frequencies, with significant low-frequency energy introduced by the sharp transients of the positional shifts. In contrast, both Regan and Beverley and Silverstein et al. used sinusoidal oscillations of a specific temporal frequency. The observed sensitivity to coherent motion at higher temporal frequencies found in the present study may well be attributable to the presence of the low-frequency components in the discrete displacements. It would be interesting to determine whether the ability to discriminate the coherent movements of our DOGs would be disrupted if those movements contained only high temporal-frequency components.

The fact that observers can discriminate motion coherence in temporally asynchronous sequences, disregarding spatial-frequency and contrast differences, is of interest because a fundamental problem in image processing is how to preserve globally coherent structural information given only local motion measurements by a set of independent and differentially selective spatial mechanisms (Yuille & Grzywacz, 1988). If the mechanisms differ in their spatial and temporal resolution, it is difficult to conceptualize a principled means for maintaining information about the global spatial structure of moving objects. It is obvious, however, that the apparent structural stability of moving objects is fundamental to our experience of the visual world, even though most objects in natural scenes produce retinal images at multiple spatial scales. Indeed, many years ago, Gestalt psychologists demonstrated that coherent motion or "common fate" was a powerful principle in organizing image segmentation.

Although the computational problem of integrating local measures into a global velocity map is often acknowledged, thorough analysis of the problem is typically neglected. Recently, however, two solutions have been proposed. Koenderink (1988) has demonstrated that, in principle, a solution to this problem can be obtained by designing mechanisms with differing time constants so as to synchronize information about the present moment by integrating over different durations of the immediate past. Stanford (1987) has shown that this may be plausible biologically in that the responses of diverse groups of ganglion cells in the cat's retina are temporally synchronized. Our results are compatible with this formalism in that performance was comparable in the three spatial-frequency conditions. Moreover, Koenderink's solution predicts that temporal resolution should decline as events grade into the past. The steady decline in the observers' discrimination accuracy with the increasing temporal delay of the right contour's displacement conforms with this prediction. Nevertheless, the ability to detect coherence in temporally asynchronous sequences does not directly follow from this formalism. On the basis of this solution, we could have quite reasonably expected complete disruption of coherent-motion perception with any temporal de-

Figure 8. Summary of results from Experiments 1 and 2 indicating discrimination accuracy in each temporal-frequency condition as a function of the duration of the delay between the relative spatial displacements of the right and left bars. Each data point represents the mean for 4 observers over five ranges of motion and three spatial frequencies (3,000 observations).

lay whatsoever. Our results show, however, that the visual system is quite tolerant of temporal delays between the signals.

Grzywacz, Smith, and Yuille (1989) developed a model that simultaneously solves for spatial and temporal coherence using a global minimization function. Computationally, coherence is achieved in two stages. In the first stage, local space-time velocity measures are computed for continuously moving points (as isolated features or as components of contours) and discrete displacements of points (apparent motion). These local measures are integrated in the second stage using a spatial smoothing operator constrained to respect local measures and a temporal operator designed to minimize changes in motion direction (motion direction is used in the temporal derivative rather than velocity because image velocity changes with depth, whereas direction stays constant—at least for translations). Again, our results are compatible with this formalism in that performance was comparable in the three spatial-frequency conditions. Nevertheless, on the basis of this formalism, we could have expected that our observers would not perceive any coherent motion whatsoever in our displays, because the temporal component of the func-

Figure 9. Summary of results from Experiments 1 and 2 for individual observers. Each data point represents the mean over five ranges of motion and three spatial frequencies (750 observations).

tion minimizes changes in motion direction. The DOGs in our displays jumped to and fro within a limited range of displacement values. In other words, motion direction was not consistent. Typically, the size and direction of the displacement changed with each displacement in both the correlated and uncorrelated displacement sequences. With motion direction (and velocity) ambiguous, the weight given a particular motion direction changed with each positional shift. Thus, the system should not be able to establish a unique, temporally smoothed solution over the global velocity field, and the coherence or incoherence of the movements should be indeterminate. When there are no temporal phase shifts in the relative displacements of the two contours, spatial smoothing alone could overcome this ambiguity in the temporal domain and account for accurate performance. Temporal delays between their relative displacements, however, should have disrupted the spatial smoothing of the velocity field. In turn, disruption in the spatial domain should have completely disrupted coherent-motion perception, but this was not found. Therefore, Grzywacz et al.'s (1989) motion-coherence model also cannot fully account for our results indicating that human observers can perceive coherent movements despite temporal delays.

It is possible that temporal integration of the local signals can account in part for the system's ability to discriminate coherence among the asynchronous signals (cf. McKee & Welch, 1985). Temporal integration could effectively suppress "noise" introduced to the system by small temporal delays between incoming signals. This fact could be incorporated successfully into the models proposed by both Koenderink (1988) and Grzywacz et al. (1989), thereby eliminating or reducing their susceptibility to temporal delays between the movements of the spatially separated features. It remains to be demonstrated, however, the extent to which temporal integration in local regions of the visual field can affect the perceived global coherence of signals originating from retinal locations separated by distances greater than a few minutes of arc.

Considering the perceptual utility of coherent-visual-motion information in image segmentation and determining object structure (Nakayama, 1985) and in directing visual attention (Driver & Baylis, 1989), it should not be surprising that the visual system can tolerate some degree of temporal delay in order to preserve the coherence of objects when they move. This would suggest, however, that any proposed mechanisms for integrating signals differing in their spatial and temporal properties should also be tolerant of temporal delays—on the order of several milliseconds—between inputs originating at noncontiguous regions of the retina.

REFERENCES

- ADELSON, E. H., & BERGEN, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America A*, **2**, 284-299.
- ADELSON, E. H., & MOVSHON, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, **300**, 523-525.
- ANDERSON, S. J., & BURR, D. C. (1985). Spatial and temporal selectivity of the human motion detection system. *Vision Research*, **8**, 1147-1154.
- DRIVER, J., & BAYLIS, G. C. (1989). Movement and visual attention: The spotlight metaphor breaks down. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 448-456.
- GEORGESON, M. A. (1987). Temporal properties of spatial contrast vision. *Vision Research*, **27**, 765-780.
- GRZYWACZ, N. M., SMITH, J. A., & YUILLE, A. L. (1989). A common theoretical framework for visual motion's spatial and temporal coherence. In *Proceedings of IEEE Workshop on Visual Motion* (pp. 148-155). Washington, DC: IEEE Computer Society Press.
- HEEGER, D. J. (1987). Model for the extraction of image flow. *Journal of the Optical Society of America A*, **4**, 1455-1471.
- KELLY, D. H. (1979). Motion and vision: 2. Stabilized spatio-temporal threshold surface. *Journal of the Optical Society of America*, **69**, 1340-1349.
- KOENDERINK, J. J. (1988). Scale-time. *Biological Cybernetics*, **58**, 159-162.
- LAPPIN, J. S., NORMAN, J. F., & MOWAFY, L. (1991). The detectability of geometric structure in rapidly changing optical patterns. *Perception*, **20**, 513-528.
- LAPPIN, J. S., WASON, T. D., & AKUTSU, H. (1987). Visual detection of common motion of spatially separate points. *Bulletin of the Psychonomic Society*, **25**, 343.
- McKEE, S. P., & WELCH, L. (1985). Sequential recruitment in the discrimination of velocity. *Journal of the Optical Society of America A*, **2**, 243-251.
- MOWAFY, L. (1991). Motion-based segregation of superimposed surfaces. *Investigative Ophthalmology & Visual Science*, **32**, 830.
- MOWAFY, L., BLAKE, R., & LAPPIN, J. S. (1990). Detection and discrimination of coherent motion. *Perception & Psychophysics*, **48**, 583-592.
- NAKAYAMA, K. (1985). Biological image motion processing: A review. *Vision Research*, **25**, 625-660.
- NAKAYAMA, K., & SILVERMAN, G. H. (1984). Temporal and spatial characteristics of the upper displacement limit for motion in random dots. *Vision Research*, **24**, 293-300.
- RAMACHANDRAN, V. S., & ANSTIS, S. M. (1983). Displacement thresholds for coherent apparent motion in random dot patterns. *Vision Research*, **12**, 1719-1724.
- REGAN, D., & BEVERLEY, K. I. (1978). Looming detectors in the human visual pathway. *Vision Research*, **18**, 415-421.
- SILVERSTEIN, D. A., KLEIN, S. A., & CARNEY, T. (1991). The detection of temporal asynchrony in two-dot targets. *Investigative Ophthalmology & Visual Science*, **32**(Suppl.), 842.
- STANFORD, L. R. (1987). Conduction velocity variations minimize conduction time differences among retinal ganglion cell axons. *Science*, **238**, 358-360.
- STONE, L. S., WATSON, A. B., & MULLIGAN, J. B. (1990). Effect of contrast on the perceived direction of a moving plaid. *Vision Research*, **30**, 1049-1067.
- VAN SANTEN, J. P. H., & SPERLING, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, **2**, 300-320.
- WATSON, A. B., & AHUMADA, A. J., JR. (1985). Model of human visual-motion sensing. *Journal of the Optical Society of America A*, **2**, 322-341.
- WATSON, A. B., THOMPSON, P. G., MURPHY, B. J., & NACHMIAS, J. (1980). Summation and discrimination of gratings moving in opposite directions. *Vision Research*, **20**, 341-347.
- WILLIAMS, D., PHILLIPS, G., & SEKULER, R. (1986). Hysteresis in the perception of motion direction as evidence for neural cooperativity. *Nature*, **324**, 253-255.
- WILLIAMS, D., & SEKULER, R. (1984). Coherent global motion percepts from stochastic local motions. *Vision Research*, **24**, 55-62.
- WILSON, H. R., & BERGEN, J. R. (1979). A four mechanism model for threshold spatial vision. *Vision Research*, **19**, 19-32.
- YUILLE, A. L., & GRZYWACZ, N. M. (1988). A computational theory for the perception of coherent motion vision. *Nature*, **333**, 71-74.

(Manuscript received August 1, 1990;

revision accepted for publication April 10, 1992.)