

## Time, Distance, and Feature Trade-Offs in Visual Apparent Motion

Peter Burt and George Sperling  
New York University and Bell Laboratories

A model of visual apparent motion is derived from four observations on path selection in ambiguous displays in which apparent motion of illuminated dots could, in principle, be perceived along many possible paths: (a) Whereas motion over each path is clearly visible when its stimulus is presented in isolation, motion is usually seen over only one path when two or more such stimuli are combined (competition). (b) Path selection is nearly independent of viewing distance (scale invariance). (c) At transition points between paths  $i$  and  $j$  (where apparent motion is equally likely to be perceived along  $i$  and  $j$ ), the time  $t$  and distance  $d$  between successive points along the paths are described by a log linear  $d/t$  relationship; that is,  $t = A - B \log (d_j/d_i)$ . (d) When successive elements along a path differ in orientation or size, the perceived motion along this path is not necessarily weaker than motion along a path composed entirely of identical elements. The model is a form of strength theory in which the path with greatest strength  $S$  becomes the dominant path. From scale invariance, we prove that the contributions of time and distance to stimulus strength are independent. From the log linear  $d/t$  relationship, we derive the precise trade-off function between  $d$  and  $t$  and show the existence of an optimal interstimulus interval to maximize the strength for any path. The model accounts well for the path-selection data and suggests a neural interpretation in which motion perception is based on the outputs of elementary detectors that are scaled replicas of each other, all having the same geometry and time delays, and differing only in size and orientation.

A visual stimulus, such as a bar or a disk, which is flashed first at one position and then flashed again nearby, may evoke a powerful illusion of movement, provided the spacing and timing of the two flashes is chosen appropriately. The vividness of this apparent motion depends strongly on the spatial and temporal separation of the stimuli and only weakly on the figural similarity of one stimulus to the other (see Kolers, 1972, for a review). However, efforts by Korte (1915), Neuhaus (1930), and others to discover a

relation between the vividness of motion and these stimulus parameters have yielded some useful heuristics but meager quantitative results.

There are two important drawbacks to these traditional studies of apparent motion. First, they are based on subjective judgments of the "quality" of perceived motion or on a subject's phenomenal descriptions of the perceived motion. Unfortunately, the "quality" of apparent motion does not correspond to a simple unidimensional perceptual continuum. In their responses, subjects weight (or confuse) a number of aspects of the motion perception, including the apparent velocity and spatial extent of the perceived motion and the presence or absence of cues that contradict motion as well as various subtle aspects of the motion "sensation" itself. For example, in the case of phi motion, produced by two adjacent rectangles, the two components of the stimulus are said to appear distinct and stationary, and yet "objectless" motion is seen between them. The two components of the stimulus are per-

---

The experimental work was carried out at New York University, where P. Burt was a postdoctoral fellow on National Eye Institute Grant EY07003. The manuscript was prepared while he was a fellow at Bell Laboratories. The authors wish to express their thanks to Eileen Kowler for serving as a subject in extensive preliminary experiments, to Geoffrey Iverson and Jean-Claude Falmagne for their advice on the proof in Appendix A, and to Geoffrey Iverson for serving as a subject in preliminary experiments and the main experiments.

Requests for reprints should be sent to P. Burt, who is now at Electrical, Computer, and Systems Engineering Department, Rensselaer Polytechnic Institute, Troy, N.Y. 12181, or to G. Sperling, Department of Psychology, New York University, New York, N.Y. 10003.

ceived as arising from different objects, and this may have an adverse influence on the judged quality of motion. Such problems suggest that quality judgments are based on a combination of perceptual and cognitive factors and are too complex to form the basis of a theory of motion perception.

A second difficulty with traditional studies is that they have almost universally employed a two-view stimulus, such as the phi inducing stimulus described above. In practical applications, such as movies and television, the stimulus for apparent motion consists of many successive exposures of the same object at spatially distinct positions—a many-view stimulus. Insofar as the two-view and many-view stimuli have been compared (Sperling, 1976), they exhibit vastly different perceptual properties. For example, the interflash interval for maximizing the judged quality of apparent motion is an order of magnitude shorter in many-view than in two-view stimuli (Sperling, 1976, Figure 5). Until the reasons for these differences between results from two-view and many-view experiments are elucidated, the many-view data are clearly more relevant to most applications.

In the present article, we describe an alternative technique for estimating the perceptual effectiveness of a motion stimulus. (This technique and preliminary experimental results have been reported in Burt, 1976, and Burt & Sperling, 1978.) The procedure relies not on a subject's cognitive judgment of quality but on a motion mechanism within the perceptual system itself that selects between alternative paths for motion. We present a multiple-path motion stimulus that is constructed in such a way that one might expect apparent motion of individual stimulus components to be seen over any of several distinct paths. These potential paths for perceived motion may differ in time, distance, and feature characteristics. Our fundamental observation is that, except at transition points, perceived motion is partially suppressed or totally absent for all but one of these competing paths. A transition point is a point on the boundary of a parameter space that separates regions of dominance of different paths. Furthermore, path dominance depends in a systematic way on the

time and distance parameters of the alternative paths.

In our technique, we exploit the path-selection phenomenon to determine the relative contributions of time, distance, and feature characteristics of the stimulus to the perception of motion. Ours is a *balance* method, in which different motion paths compete against each other like weights in a balance, but only one path achieves perceptual dominance. In our first experiment, we demonstrated and measured the path-selection phenomenon.

In Experiment 2, we investigated the time-distance trade-off in apparent motion. Korte's second law (Korte, 1915) proposes a direct relation between the optimal temporal interval for apparent motion and the spatial separation of the stimulus components. On the contrary, we find scale invariance: Path selection does not depend on the overall scale of the stimulus (i.e., path selection is independent of viewing distance, which means that the optimal interval does not vary with spatial separation). Scale invariance is a critical property because it implies independent contributions of time and distance to the selection process.

In the third experiment, we investigated the sensitivity of apparent motion to figural agreement between stimulus components. Is there a preference for motion between like components, such as identically oriented short line segments, over motion between unlike components, such as orthogonally oriented line segments? We find no measurable preference for motion between figurally similar elements over dissimilar elements.

When we make the assumption that the dominant path is the one containing the strongest motion stimulus, we account for the observed pattern of path dominance with a simple quantitative relationship between the strength measure, which we define for the motion stimulus, and the time, distance, and feature characteristics of the stimulus' components. Within the range of time and distance intervals examined, we find that there is an optimal time interval of about 20 msec that maximizes strength of apparent motion, and there is a monotonic inverse relation between stimulus strength and distance.

We then show that our mathematical theory is consistent with a neural model in which the initial processing of motion is performed by elementary *motion detecting units*. Individual units perform their analysis in small, local regions. Motion is inferred when a correlation occurs between stimulus events within one subregion and earlier (delayed) events in an adjacent subregion. The subareas are insensitive to image detail such as contour orientation. Although it is possible to make more complex models, the simplest neural model to account for our data assumes that all units have the same temporal delay properties and differ spatially only in magnification and rotation, that is, they are all scaled replicas of a canonical unit with, possibly, an angular change of orientation with respect to the horizontal. The experiments to be reported do not offer a strong proof of this model, but we offer it to assist the reader in interpreting the data and the theory.

*The Multiple-Path Motion Stimulus*

The multiple-path motion stimulus consists of a single horizontal row of dots that is flashed very briefly on a CRT screen (see Figure 1, Panel a). The dots within the row are evenly spaced with a separation distance  $D$ . The row itself is presented again and again, with a time interval  $t$  between successive presentations. Between each two presentations, the row is displaced horizontally on the screen by a distance  $H$  and vertically (downward) by a distance  $V$ .

Each horizontal row of dots in Figure 1, Panel a represents the position of the stimulus row at a different moment in time. If at time  $T_0$  the row is at the position indicated by the filled dots at the top of the figure, then at time  $T_1 = T_0 + t$  the row is presented at the position of the top row of open dots. With repeated presentations the row steps down the screen and to the right. New dots are introduced at the left end of the row as dots pass off the screen at the right.

The unique property of this dot configuration is that it is a stimulus for motion over several different paths at once. Three of these competing paths are indicated by arrows in Figure 1, Panel a. The path marked

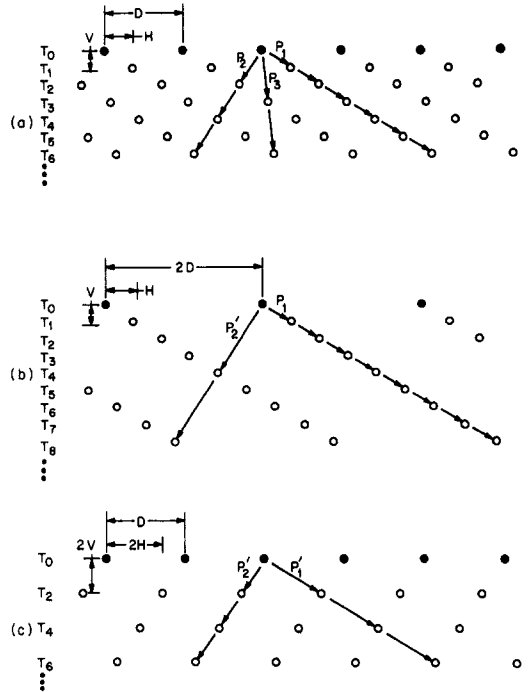


Figure 1. Ambiguous motion stimuli. (Panel a shows a multiple-path motion stimulus  $M_{1,2}$  generated by repeatedly flashing a horizontally oriented row of dots on a CRT screen. Dot spacing within the row is  $D$ . With each new presentation the row is displaced downward a distance  $V$  and to the right a distance  $H$ . Solid circles show the position of dots at time  $T_0$ ; open circles show dot positions at subsequent times  $T_i$ , where  $T_i = T_0 + it$ . Arrows show some possible paths for apparent motion of a dot presented at time  $T_0$ . Path  $P_1$  represents apparent motion to the position of the nearest dot at time  $T_i$ . Generally, all dots of the row appear to move together along the same path. Path dominance is determined by the particular values of  $t$ ,  $D$ ,  $V$  and  $H$ . Panel b shows stimulus  $M_1$ , which contains a subset of the dots of stimulus  $M_{1,2}$ : Every other dot has been removed. Path  $P_1$  is unchanged, whereas,  $P_2'$  and higher number paths are greatly altered. Panel c shows stimulus  $M_2$ , which contains another subset of  $M_{1,2}$ : Every other row has been removed. Path  $P_2'$  is unchanged, but the distance between dots along path  $P_1'$  has been doubled relative to  $P_1$  in  $M_{1,2}$ .  $P_1'$  and  $P_2'$  in  $M_1$  and  $M_2$  have the same velocity and direction as  $P_1$  and  $P_2$  in  $M_{1,2}$ ; they differ in dot density along the path.)

$P_1$  shows a dot in the row presented at time  $T_0$  moving to the position of the nearest dot in the next row presentation. Path  $P_2$  shows the same initial dot moving to the position of the nearest dot in the next row but one. In general,  $P_n$  is the path to the nearest dot in the row presented at  $T_n$ ,  $n$  time intervals later. The time interval between dots in path

$P_n$  is  $nt$ , where  $t$  is the interval between successive row presentations. In principle, a dot in one row presentation could appear to move to the position of any dot in a subsequent row presentation. We give names only to paths between nearest dots because apparent motion was observed only along these paths.

Although the dot configuration of Figure 1, Panel a, contains a stimulus for motion over many distinct paths simultaneously, we have found that an observer will generally see only one of these motions. Path dominance is determined by the particular values assigned to the display parameters  $D$ ,  $V$ ,  $H$ , and  $t$ . (The subject must carefully fixate a stationary point on the screen while making these observations, since eye tracking in the direction of any path greatly enhances the perception of motion along that particular path.) Stimuli characterized by a single geometric configuration (i.e., particular  $D$ ,  $H$ , and  $V$ ) can evoke dominant motion over any of several different paths, depending on the value of the interflash interval,  $t$ . In this case, there are distinct ranges of  $t$  that favor each motion path.

When motion along a particular path is suppressed, conscious effort cannot make motion along this path re-appear. Taken together with our other observations, this leads us to suggest that there is a preconscious perceptual decision process involved in motion perception and that in reaching its decision, the process weighs both time and distance characteristics of the competing paths. This idea is supported by a further observation: When the dot configuration is modified to isolate the stimulus for motion over any particular path (by removing competing paths), then that motion may be clearly visible even for values of  $t$  at which motion along this path is completely suppressed in a combined stimulus.

### Experiment 1

In this experiment we characterize the perceptual path selection phenomena by measuring the degree to which motion sensations for one path are suppressed by the presence of stimuli for competing paths. We used a subjective judgment to calibrate the balance method as follows: First, subjects

viewed motions on paths  $P_1$  and  $P_2$  in the multiple-path motion stimulus and compared them to the appearance of these motions when presented in isolation. Then, the subjects rated the strength of motion in the combined stimulus as a fraction of its strength when seen in isolation. The subjective strength measure used here is not to be confused with the strength function inferred from the results of later experiments. (The two strength measures are consistent insofar as they may be compared.)

### Method

*Stimulus.* Three types of stimuli were presented. The first,  $M_{1,2}$  is the multiple-path motion stimulus as shown in Figure 1, Panel a, in which  $D = 1.70$  cm,  $H = \pm .68$  cm, and  $V = .18$  cm. These distances represent small whole numbers of raster units on our screen. At the viewing distance of 2 m,  $D$ ,  $V$ , and  $H$  subtend, respectively, 29.3, 11.7, and 3.09 minutes of visual angle. This assignment of values to  $D$ ,  $V$ , and  $H$  was chosen because, in preliminary tests, we found that the stimulus could evoke either  $P_1$  or  $P_2$  motions, depending on  $t$ . Also, in this stimulus,  $P_1$  and  $P_2$  are in opposite directions with respect to the vertical midline: When  $H$  is positive,  $P_1$  is a rightward motion and  $P_2$  is a leftward motion; when  $H$  is negative, these directions are reversed.

The other two stimulus configurations,  $M_1$  and  $M_2$ , are shown in Figure 1, Panels b and c. Stimulus  $M_1$  is so called because it is constructed from  $M_{1,2}$  in a way that selectively alters the stimulus for path  $P_2$  and thereby favors motion along  $P_1$ . This is done by eliminating every other dot in each row of the presentation, or, equivalently, by doubling  $D$ . Along path  $P_1$ , the time and distance intervals between successive dots are identical in  $M_1$  and in  $M_{1,2}$ ; along path  $P_2$  these distances are doubled in  $M_1$  relative to  $M_{1,2}$ . The effect of this doubling is to greatly impair  $P_2$  as a candidate perception in  $M_1$ , so that  $P_1$  is seen in virtual isolation.

Stimulus  $M_2$ , on the other hand, presents motion along path  $P_2$  in relative isolation. It is constructed from  $M_{1,2}$  by eliminating every other row. This is equivalent to doubling  $H$ ,  $V$ , and  $t$ . Along path  $P_2$ , the time

and distance separations of dots are identical in  $M_{1,2}$  and in  $M_2$ ; along path  $P_1$ , these intervals are twice as large in stimulus  $M_2$  as in  $M_{1,2}$ . Thus, the stimulus for motion along  $P_1$  is exceedingly weak in  $M_2$  compared to  $M_{1,2}$ , but the stimulus for motion along path  $P_2$  has the same time and distance parameters in both configurations. The stimulus parameters are summarized in Table 1.

**Display.** Subjects sat 2 m from the display in a room that was dark except for the stimulus itself and two incandescent lamps that illuminated the display surface. The display was viewed binocularly without head restraint. The dot stimulus was generated by a Digital Equipment Corporation PDP-15 computer on a VT/15 oscilloscope display with a P4, fast, white phosphor. The background luminance of the display surface due to room lighting was  $.058 \text{ cd/m}^2$ , and the dots drawn on the screen had a net luminous directional energy of about  $1.6 \times 10^{-6}$  candelas per refresh (Sperling, 1971). Dots were refreshed twice with each row presentation to increase their net energy; the second refresh followed the first by 3 msec. The dot rows appeared within a rectangular area 22 cm long (6.3 degrees when viewed at 2 m) and 6 cm (1.72 degrees) high; the bottom of this area was 4 cm (1.15 degrees) above the fixation point.

**Procedure.** The experiment was run as a series of trials. In each trial, an ambiguous  $M_{1,2}$  and one of the two related stimuli  $M_1$ ,  $M_2$ , were presented alternately.  $M_1$  and  $M_2$  served as control (reference) stimuli on which to base judgments of the strength of motion in  $M_{1,2}$ . On half of the trials, the subject was required to judge the strength of  $P_1$ . On these trials,  $M_1$  and  $M_{1,2}$  were alternated three times, beginning with  $M_1$ . There was a 1-sec pause between successive stimuli.

Subjects were required to estimate the "strength" of  $P_1$  in  $M_{1,2}$  relative to  $M_1$ . At the end of the trial, they reported this estimate on a scale of 0-6: 0 indicated that  $P_1$  was undetectable, 1 that it was 20% as strong as the reference, and so on up to 5, which indicated that  $P_1$  appeared of equal strength in both stimuli. A judgment of 6 indicated that motion appeared stronger in  $M_{1,2}$  than in  $M_1$ .

Table 1  
Stimulus Configurations for Experiment 1

Measure	Configuration		
	$M_{1,2}$	$M_1$	$M_2$
Dot separation in row	$D(17 \text{ mm})$	$2D$	$D$
Horizontal row displacement	$H(6.8 \text{ mm})$	$H$	$2H$
Vertical row displacement	$V(1.8 \text{ mm})$	$V$	$2V$
Time between presentations	$t(12-56 \text{ msec})$	$t$	$2t$
Dominant path <sup>a</sup>	$P_1$ or $P_2$	$P_1$	$P_2$
Viewing distance	2 m		

<sup>a</sup> Path  $P_2$  in  $M_2$  is equivalent to  $P_2$  in  $M_{1,2}$ .

In separate trials  $M_2$  was alternated with  $M_{1,2}$ , in order to estimate the strength of motion along path  $P_2$ . On half of the trials, the left to right mirror images of the stimuli shown in Figure 1 were presented to obtain motions in opposite directions. Thus, there were four kinds of trials run in a pseudo-random order:  $P_1$  right,  $P_1$  left,  $P_2$  right,  $P_2$  left. Ten values of  $t$  were used on different trials: These ranged in logarithmic steps from 12 to 56 msec. Within a trial, the same  $t$  was used for the ambiguous  $M_{1,2}$  stimulus and the control, so time and distance parameters for the two apparent motions being compared on the trial were identical.

Two subjects (the authors) served in this experiment. Forty trials were run in each session of the experiment, one for each combination of the four motion types and 10 values of  $t$ . There were four such sessions on four different days.

## Results

The results of Experiment 1 are shown in Figure 2. Each data point represents the average of four observations. Consider first the judgments made by subject PB shown in Figure 2, Panel a. Here,  $P_1$  motion is to the right and  $P_2$  motion is to the left. We see that when  $t$  was greater than 34 msec, the  $P_1$  motion was judged only slightly less strong in the multiple-path motion stimulus  $M_{1,2}$  than in the isolated stimulus  $M_1$ . Similarly, for  $t$  less than 17 msec,  $P_2$  is judged at least 80% as strong in the ambiguous  $M_{1,2}$  stimulus as in its control stimulus  $M_2$ . On the

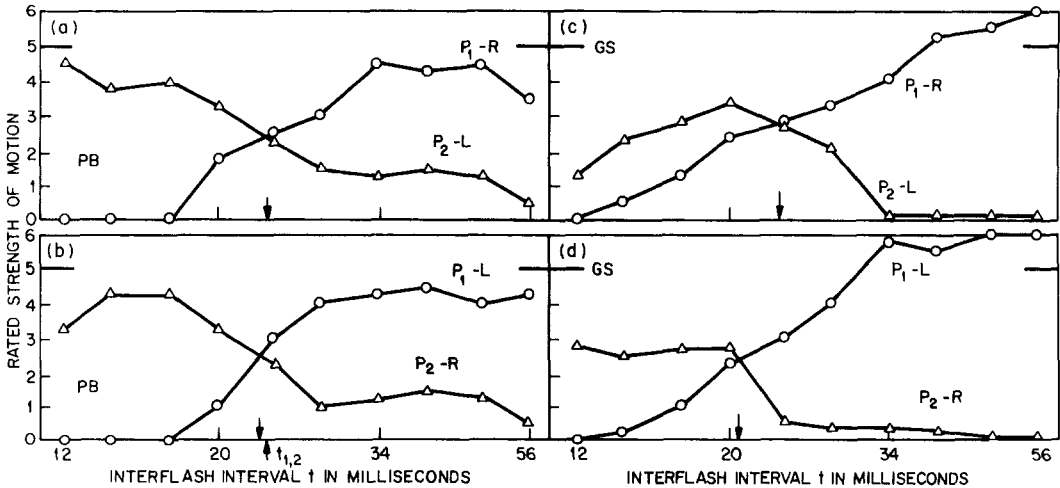


Figure 2. Strength of apparent motion in an ambiguous stimulus compared to control stimuli. (Data are shown for two subjects, PB and GS. The interflash intervals  $t$  of the ambiguous and of the control stimuli are indicated on the abscissa in a log scale. The ordinate represents the judged strength of apparent motion on paths  $P_1$  and  $P_2$  in the ambiguous stimulus  $M_{1,2}$  [Figure 1, Panel a] compared with its strength when presented in isolation in the comparison stimuli  $M_1$  [Figure 1, Panel b] and in  $M_2$  [Figure 1, Panel c]. L = left and R = right directions of motion. Motions judged to be as strong in  $M_{1,2}$  as in  $M_1$  or  $M_2$  were given a rating of 5; higher and lower ratings for either path indicate proportionate increases or reductions in motion strength in the combined stimulus. Observations were made at 10 different values of  $t$  for a single spatial configuration [Panels a and c] and for the left-to-right reflection of this configuration [Panels b and d]. The arrows pointing downward indicate the crossovers in Experiment 1; the transition value  $t_{1,2}$  measured for PB in Experiment 2 is indicated by the arrow pointing upward under the abscissa of Panel b.)

other hand, for the three smallest  $t$ s, apparent motion along path  $P_1$  was completely suppressed in  $M_{1,2}$ ; it was not detected on a single one of 12 trials. Similarly, motion  $P_2$  in  $M_{1,2}$  was severely reduced for large  $t$  although it was not completely suppressed.

These results demonstrate regions of selective suppression, regions in which apparent motion along one path is relatively unaffected by the presence of the stimulus for a second path, whereas the apparent motion along the second path is severely reduced by the first. There is a transition zone between  $t$  values of 20 and 30 msec in which dominance gradually shifts from  $P_2$  to  $P_1$ . The strength curves cross when  $t$  equals about 23 msec, and at this point apparent motion along each path in the ambiguous stimulus was judged to be about 50% of its strength in the control stimulus. The particular dot configuration used in this experiment was chosen because it had large regions of  $P_1$  and  $P_2$  dominance. Apparent motion along other

paths ( $P_3, P_4$ , etc.) became dominant when  $t$  was less than 12 msec; these paths were not the object of study here and therefore we did not explore smaller values of  $t$ .

We should also emphasize that apparent motion was clearly seen in all control stimuli throughout the range of time values used in this experiment; that is, strong apparent motion was seen in control stimuli even for values of  $t$  at which motion along this path was greatly or totally suppressed in the ambiguous motion stimulus. For very large  $t$ s, greater than about 50 msec, the quality of apparent motion would deteriorate even in control stimuli; the stimulus would be seen as sequentially flashed rows with little apparent motion between them.

For a range of large  $t$ s that produce good apparent motion, the stimulus appears as a single row moving down the screen. In this case, the apparent spacing of the dots in the row depends on the dominant motion path: For a path  $P_n$  the apparent spacing is  $D/n$ .

The geometric basis of this illusion is obvious from Figure 1.

For very small  $t$  values, motion quality is again degraded. In this case, one perceives not one row in motion but a group of several moving together. For  $t$ s less than about 10 msec, the number of rows that seem to be simultaneously visible can become so large that the display fills much of the display area. These observations may be related to the regions of simultaneity and succession described by numerous previous investigators (Kolars, 1972), and the grouping of stimuli has been reported by Ross (Note 1).

If we now compare the results shown in Figure 2, Panel a, with those in Panels b, c, and d, we see some systematic differences. Subject GS reported motion  $P_1$  as being somewhat stronger in the ambiguous stimulus over the entire  $t$  range than did subject PB, and he rated  $P_2$  as less strong. This difference in the data may reflect a bias in use of the subjective ratings or it may reflect a real difference in the subjects' relative sensitivity to motions along these paths. There were certainly real differences at small values of  $t$ , where subject GS reported that higher order apparent motions along paths  $P_3, P_4$ , and so forth interfered with his ability to perceive motion along path  $P_2$ . It may be that small differences in sensitivity will have large effects on motion dominance with the particular stimulus configuration used in this experiment, especially at small values of  $t$ . (This possibility is supported by Experiment 2 and by the subsequent analysis.)

The transition from seeing apparent motion along path  $P_2$  to  $P_1$  occurred at nearly the same value of  $t$  for both subjects. At this transition point, both subjects rated the strength of motion along each path in the ambiguous stimulus at about 50% of its strength in the controls, which suggests that they used the rating scale in a highly consistent manner.

## Experiment 2

Experiment 1 demonstrated that the perception of apparent motion along some paths was selectively suppressed when a stimulus for apparent motion on a competing path

was present. In Experiment 2, we used the path selection phenomena to discover the relative contribution of time and distance to the perception of apparent motion.

With the ambiguous multiple-path motion stimulus of Experiment 1, different motion paths dominated perception for different presentation rates. This demonstrates that the selection process is sensitive to the time intervals between stimulus points along competing paths. Since in Experiment 1 we used a single stimulus configuration, nothing could be discovered about the possible contribution of dot separation to selection. However, from similar, but informal, experiments with other configurations, we found that  $t$  regions of dominance always occurred, but transitions between dominant paths did not always occur at the same value of  $t$ . This observation indicates that the selective suppression mechanism is sensitive to the spatial parameters of the stimuli.

In Experiment 2, we determined the influence of dot separation on motion selection by measuring the  $t_{1,n}$  transition points for a large number of different dot configurations. These transition points are the  $t$  values at which dominance shifts from Motion Path 1 to Motion Path  $n$ . In graphs, such as those in Figure 2, the  $t_{1,2}$  transition point is the  $t$  value at which the strengths of paths  $P_1$  and  $P_2$  cross. This single  $t_{1,2}$  transition point is a highly efficient characterization of an extensive set of data. Little of importance is lost by focusing on transition points, and it makes practical the investigation of a much larger number of conditions than would otherwise be possible.

Of the many possible outcomes of this experiment, we shall consider two very different possibilities. Korte's laws (1915), based on his classical two-view experiments, predict velocity invariance: The optimal time between dots is proportional to the retinal distance between them.<sup>1</sup> If this were the case in our experiment, then the transition times

<sup>1</sup> To derive predictions of our experiment from Korte's second law requires additional assumptions. We assume the ordering of motion strengths as a function of velocity does not depend on the spatial distance between successive points along the motion path.

obtained at 1 m would be twice as large as those obtained at 2 m.

The antithesis of velocity invariance is scale invariance: Changes in stimulus scale (by varying viewing distance) have little or no effect on the transition time. Scale invariance is a particularly significant property because it indicates that the motion selection mechanism is primarily sensitive to relative, not absolute, distances between dots along competing paths. Both velocity invariance and scale invariance—insofar as they hold—have profound implications for the underlying mechanism of motion perception.

### Method

*Subjects.* Two subjects participated in the experiment, including one of the authors. Both were experienced observers in motion experiments.

*Stimuli.* Eighteen multiple-motion dot configurations were generated on a CRT

screen (see Table 2). Each of these was viewed from both 1 and 2 m. (Changing the viewing distance alters the scale of the configurations, that is, the absolute distances between dots, without changing the relative distances.) One transition point was measured for each configuration. Of these measured transition points, one third were for transitions between paths  $P_1$  and  $P_2$ , one third for  $P_1$  and  $P_3$ , and the remainder for  $P_1$  and  $P_4$ .

Two criteria were used in choosing dot configurations for this experiment: (a) for  $t$  near that at which the transition of interest occurred, only motion along two competing paths should be visible; and (b) one of these motions should be to the right and the other should be to the left. These constraints were intended to minimize any possible confusion in identifying the motion paths.

A stimulus presentation consisted of a single sweep of the dot row through the rectangular display area defined in Experiment

Table 2  
*Stimulus Configurations and Results for Subjects PB and GI in Experiment 2*

Transition type	Parameter			Viewing distance			
				1 m		2 m	
	$H$	$V$	$d_n/d_1$	Transition time		Transition time	
			PB	GI	PB	GI	
$P_1-P_2$	6.81	1.84	.71	24.5	31.7	23.5	29.0
		2.45	.82	17.5	24.0	19.0	25.3
		3.06	.93	15.0	22.0	17.0	24.5
	7.23	1.84	.60	32.0	39.5	29.0	36.0
		2.45	.72	26.5	32.5	23.0	32.5
		3.06	.84	21.5	27.0	21.5	30.0
$P_1-P_3$	5.11	.82	.57	20.0	24.5	16.5	19.7
		1.22	.76	14.5	17.3	13.5	16.7
		1.63	.96	9.5	12.7	11.5	15.0
	5.32	.82	.49	24.0	27.0	18.5	21.3
		1.22	.69	17.5	20.3	15.0	18.7
		1.63	.89	13.0	16.5	13.0	16.0
$P_1-P_4$	3.83	.41	.61	11.5	13.5	10.0	11.0
		.61	.76	9.5	10.3	9.5	10.7
		.82	.93	7.8	8.5	8.3	9.8
	4.04	.41	.45	14.0	18.0	12.3	13.3
		.61	.62	11.5	16.0	10.3	12.3
		.82	.81	9.3	12.0	9.8	11.0

Note. Observed transition times are given in msec;  $H$  and  $V$  are in mm;  $D = 17$  mm (see Figure 1, Panel a).



1. The slowest moving displays were terminated after 3 sec whether or not the dot row had reached the bottom of the display area.

*Procedure: Constant stimuli, forced-choice response.* In order to find the transition point for a given dot configuration, that configuration was presented at seven different  $t$  values. These  $t$ s covered a range that was known from preliminary experiments to include the transition. After each of these presentations, the subject reported whether dominant motion was toward the right or the left. Thus, for the largest  $t$ , the subject reliably reported that dominant motion was in the direction corresponding to path  $P_1$ , whereas at the smallest  $t$ , dominant motion was reported to be in the other direction, corresponding to  $P_2$ ,  $P_3$ , or  $P_4$ .

A total of eight responses were gathered for each  $t$  value within a session of the experiment. Responses were forced choices between right and left. To cancel any directional response bias, four responses were obtained with the dot configuration as defined in Table 2, and four were obtained for its left-to-right mirror reflection. These eight responses were averaged to obtain an estimate of the subject's preference for motion  $P_1$  (responses for left and right directed  $P_1$  motion were averaged together). Except for occasional fluctuations, the fraction of  $P_1$  responses (the psychometric function) increased monotonically with  $t$ . A smooth curve was fitted by eye to the psychometric function, and the  $t$  value corresponding to the point at which the fitted curve crossed 50% was taken as the  $t_{1,n}$  transition value between paths  $P_1$  and  $P_n$ . It corresponds to the  $t$  at which the subject would have reported dominant motion over each of the two paths with equal probability.

*Sessions.* Within each session of the experiment, a subject viewed the display from one of the two distances and observed dot configurations that produced a single type of transition, for example,  $P_1$  to  $P_2$  transitions.

A total of 84 distinct stimuli were presented within a session (6 dot configurations and their reflections, each at 7 values of  $t$ ). These distinct stimuli were presented in random order and all were repeated four times. A session lasted about 30 minutes, so three sessions normally were run on a given day.

Sessions for the three transition types and two viewing distances were run in a balanced order.

### Results

Transition points for the two subjects are listed in Table 2. Each of these values is based on 168 trials; it is the mean of three determinations obtained in separate experimental sessions. The standard deviations of these three determinations of  $t_{1,n}$  were computed separately for the various ranges of  $t$ . For  $t_{1,n}$  between 0 and 10 msec, the average standard deviation is  $\sigma = .4$  msec; for  $t_{1,n}$  between 10 and 20,  $\sigma = .6$ ; for  $t_{1,n}$  between 20 and 30,  $\sigma = 1.8$ ; for  $t_{1,n}$  between 30 and 40,  $\sigma = 4.0$ . The standard errors of the mean  $t_{1,n}$  are smaller by a factor of  $1/\sqrt{2}$ .

Two aspects of these data should be noted: Although the transition points are distributed over a wide range of time values  $t$ , those for the  $P_1$  to  $P_2$  transitions occupy the high end of this range, and those for  $P_1$  to  $P_4$  the low end. On the other hand, the transition points obtained for equivalent physical stimuli at the two viewing distances differ very little; on the average,  $t$ s for the 1-m observations exceed those for 2 m by only 6%.

*Scale invariance.* The stability of transition points as viewing distance is altered implies scale invariance of the path selection process. We believe that scale invariance in motion path selection has not been reported previously. Scale invariance does not hold perfectly, particularly for  $P_1$ - $P_3$  transitions, but it holds very well. This result is in gross contrast to Korte's classic assertion of velocity invariance, which would require a 100% change in  $t$  between the two viewing distances in contrast to the 6% we found. To the extent that scale invariance holds and other factors (such as the angle between paths) can be ignored, only relative distance matters. In cases for which this is true, it is appropriate to examine the dependence of transition points on the ratio  $d_i/d_j$ , where  $d_i$  and  $d_j$  are the distances between successive dots presented along paths  $P_i$  and  $P_j$ , respectively.

*Linear  $t_{1,n}$  versus  $\log(d_n/d_1)$ .* Transition data are plotted as a function of  $\log(d_n/d_1)$  in Figure 3. (The solid lines in this figure represent predicted transition times accord-

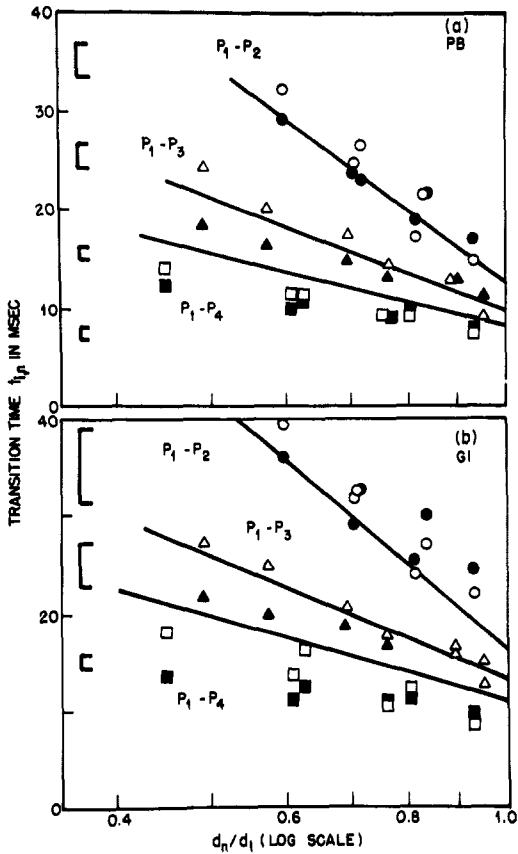


Figure 3. Transition times  $t_{1,n}$  for the 36 stimulus configurations listed in Table 1. (Data are shown for two subjects, PB and GI. These include three transition types,  $P_1$  to  $P_2$  [circles],  $P_1$  to  $P_3$  [triangles], and  $P_1$  to  $P_4$  [squares] and two viewing distances, 2 m [closed symbols] and 1 m [open symbols]. Data for two subjects are shown separately in Panels a and b. The abscissa indicates the relative distance  $d_n/d_1$ , between adjacent points on competing paths, and the ordinate indicates the transition time. Each  $t_{1,n}$  was determined from data obtained in three sessions, or a total of 168 trials. The vertical bars at the far left represent two standard deviations of the data in each of four ranges of obtained data:  $t_{1,n} = [0,10], [10,20], [20,30], [30,40]$ . See text for details. The straight lines show theoretical transition times derived from the first model developed in the Discussion section [Equation 10].)

ing to a simple model, which will be described in the Discussion section.) This graphic representation reveals a striking additional property of the data. The  $t_{1,n}$  transition values decrease approximately linearly as the log ( $d_n/d_1$ ) is increased. The scatter of the  $t_{1,n}$  data points for a particular value of  $d_n/d_1$  is due mainly to the previously men-

tioned effect of viewing distance (imperfect scale invariance). There is a small but statistically significant residual scatter in observed  $t_{1,n}$  when geometric configurations that differ in the magnitude of  $d_1, d_n$  happen to produce the same or nearly the same  $d_n/d_1$ . A part of this residual effect should be attributable to different absolute sizes of the component  $d_1$  and  $d_n$ , even at the same viewing distance. However, it appears likely that most of the residual scatter is due to the unintended presence of apparent motion along other paths (e.g.,  $P_{n+1}$  or  $P_{n-1}$ ) in particular configurations, and the influence of this third motion path on the  $P_1-P_n$  transition under study.

*Optimal interdot interval.* Although none of our observations were for values of  $d_n/d_1$  quite as large as unity, where stimulus separations over competing paths are identical, a small extrapolation to this value indicates that for the three transition types, transitions would occur there, with  $t$  values between 8 and 20 msec, depending on  $P_n$ . Thus, when the distances between points on competing paths  $P_1$  and  $P_n$ ,  $n = 2, 3, 4$  are identical, there will be a range of  $t$  for which dominant motion will be over the path  $P_n$ . This means that paths with a larger time interval between point presentations ( $nt$  for paths  $P_n$ ) can compete successfully against a path with a shorter time interval ( $t$  for path  $P_1$ ) even when the interpoint distances are the same on the competing paths. From this, the reader may discern what we prove formally later: that between  $t_{1,n}$  and  $nt_{1,n}$  there is an optimal interdot time interval for apparent motion and that  $t_{1,n}$  and  $nt_{1,n}$  are symmetrically balanced on each side of this optimum.

*Other methods and subjects.* In extensive preliminary experiments, the subjects of Experiment 2 used a method of adjustment (rather than the method of constant stimuli) and obtained quite similar results. Although we have informal data from several other observers that conform closely to the data reported here, we obtained significantly different data from one subject (EK) by the method of adjustment. EK had great difficulty making judgments of  $P_1$  to  $P_3$  and  $P_1$  to  $P_4$  transitions. Her  $P_1$  to  $P_2$  transitions conformed to scale invariance and the log

linear  $t$  versus  $d_2/d_1$  trade-off as did those of our other subjects. However her transitions were larger than those of other subjects by a factor of about 1.5. The range of individual variation and its causes are matters that require further study.

### Experiment 3

Motion cannot be detected instantaneously; detection ultimately must rely on a matching process in which a stimulus pattern seen at one moment in time is matched with a similar but displaced pattern seen at a slightly later moment in time. In Experiment 2, we examined time-distance trade-offs in this matching process. In Experiment 3, we wished to determine what types of image patterns are matched. Since the strongest motion perception occurs over short dot distances (see Discussion), we anticipate that the patterns that are matched have a correspondingly small size. We consider here the local patterns to which motion perception is sensitive.

Two critical questions about these local patterns concern (a) their complexity: What level of detail does the motion system resolve in individual image elements? and (b) the similarity of successively presented patterns: How similar must these patterns be in order to be matched? These two characteristics of local patterns determine the image elements to which the motion system will respond and the ability of the system to discriminate appropriate from false matches. We investigated these issues by measuring the strength of motion along paths composed entirely of similar elements when competing paths existed that were composed of mixtures of dissimilar elements.

### Method

*Stimuli.* Multiple-path motion stimuli were constructed as before except that the dots were replaced by other elements, such as line segments. In some cases, two different kinds of elements were alternated within a row. For example, Figure 4, Panel a, shows a stimulus configuration in which two line elements of different orientation are alternated within each row. Path  $P_1$  is composed

entirely of similar elements, path  $P_2$  is composed of alternating dissimilar elements. If motion perception is sensitive to the shape of these local elements, so that apparent motion along paths composed of similar elements is preferred, then element shape in this stimulus should bias motion dominance in favor of paths  $P_1$ . This dominance should extend over the whole range of  $t$  values for which  $P_1$  dominated in the previous experiment plus some additional values for which  $P_2$  was preferred in the previous experiment. A relative strengthening of apparent motion along paths composed of similar elements is reflected, in this experiment, by a shift in the  $t_{1,2}$  transition point to smaller  $t$ .

Stimuli composed of four different combinations of elements were used in Experiment 3: two kinds in which dissimilar elements were alternated within each row (experimental conditions) and two kinds that were constructed of a single element type (control conditions). The two alternating kinds are shown in Figure 4, Panels a and b; one is constructed from orthogonal line elements and the other from dots and rings. A third kind of stimulus was composed only of dots and was identical to the stimulus used in Experiment 2. The fourth kind of stimulus was composed entirely of rings. Control conditions with only dots or only rings were run along with the test conditions, since the critical measure was the shift in transition time between the alternated (test) and nonalternated (control) condition. Details of the stimulus elements are shown in Figure 4, Panel c.

*Procedure.* The  $P_1$ - $P_2$  transition points  $t_{1,2}$  were determined for stimuli characterized by three geometric configurations and the four combinations of elements described above (see Table 3). The same psychophysical procedure and subjects were used as in Experiment 2. A separate block of 84 trials was run for each of the four element combinations. These included two repetitions of trials for each of the three geometric configurations and their left-right reflections, and seven  $t$  values. When a block of trials had been completed for each of the element combinations, the four blocks were repeated in reverse order. A separate determination was made of the transition point  $t_{1,2}$  in each

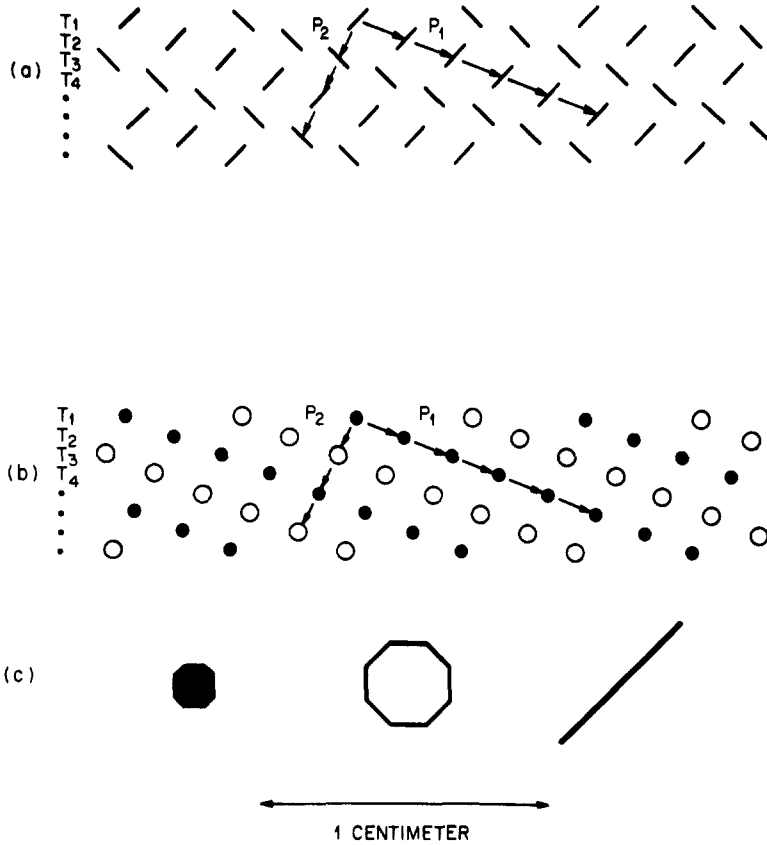


Figure 4. Stimuli comparing motion strength in paths composed of like and unlike elements. (Stimulus parameters  $D$ ,  $V$ , and  $H$  in Panels a and b are the same as in Figure 1, Panel a, except that two different elements are alternated within each row. As a result, motion on path  $P_1$  is between like elements, whereas motion on path  $P_2$  is between unlike elements. Panel c shows greatly enlarged details of individual elements. The luminous directional energies of the elements were: dot,  $1.6 \times 10^{-6}$ ; ring,  $1.9 \times 10^{-6}$ ; bar,  $10.3 \times 10^{-6}$  candela-sec per refresh.)

block for each class of stimulus (geometric configuration  $\times$  element combination).

### Results

The average of the two determinations of each transition point is given in Table 3 and illustrated in Figure 5. Each point is based on 56 trials.

*Variation of  $t_{1,2}$  with element composition.* The critical question of Experiment 3 is, What is the effect of element composition of the stimuli on  $t_{1,2}$ ? The dot-dot transition data for both subjects in Experiment 3 are very near the data obtained under equivalent conditions in Experiment 2. These data may be used as a standard against

which to measure the influence of other feature combinations on motion perception. For both subjects, the ring-ring stimulus produced transitions at very nearly the same  $t_{1,2}$  as did the dot-dot reference. Thus, if the element shape influences motion sensitivity in this case, its influence must be the same for motion on both paths.

For subject PB, the data for all four element types are remarkably similar, indicating that element composition has almost no effect on  $t_{1,2}$ . The  $t_{1,2}$  data of subject GI show a larger variation with element composition. Although the effect of alternating dots and rings is to slightly favor the homogeneous motion path,  $P_1$  ( $t_{1,2}$  for the dot-ring stimulus averages 4 msec less than the  $t_{1,2}$  of the

control), the effect of alternating the orientation of line segments is to favor the heterogeneous motion path,  $P_2$  ( $t_{1,2}$  for the lines averages 2 msec greater than that of the control). Thus alternation of elements had a very small effect for subject PB and a small but inconsistent effect for subject GI. The shift for GI cannot reflect any preference for motion between like elements, that is, along path  $P_1$ , since the shift to a larger  $t_{1,2}$  indicates a relative strengthening of heterogeneous paths.

*Energy versus pattern.* Finally, data for the dot-ring stimuli do show a decrease for both subjects, suggesting that in this case, at least, subjects are more sensitive to motion between like patterns. However another explanation should be considered: that matches are made between elements that have similar luminous energy rather than similar pattern. Here the rings have a slightly larger total energy than the dots (see caption to Figure 4). This explanation was supported by an

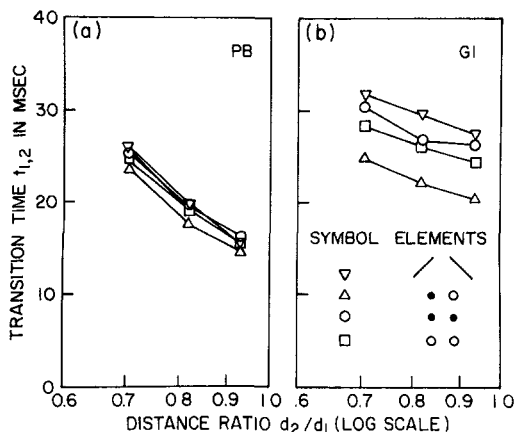


Figure 5. Transition times for motion between like and unlike elements. (Transition times are shown on the ordinate; the relative distances between elements on paths  $P_1$  and  $P_2$  are shown on the abscissa. Data were obtained for four element combinations: dots only [circles], rings only [squares], dots and rings alternated [triangles], and orthogonal line elements alternated [inverted triangles]. Each element combination was presented in the three configurations listed in Table 1. Panels a and b show data of subjects PB and GI, respectively.)

Table 3  
Stimulus Configurations and Results for Experiment 3

Elements	V	Transition time	
		PB	GI
Dots only (control conditions)	1.84	25.0	30.0
	2.45	19.5	26.5
	3.06	16.0	26.0
Mean		20.2	27.5
Rings only (control conditions)	1.84	24.5	28.0
	2.45	19.0	26.0
	3.06	15.5	24.0
Mean		19.7	26.0
Dots and rings	1.84	23.5	24.5
	2.45	17.5	22.0
	3.06	14.5	20.0
Mean		18.5	22.2
Orthogonal line elements	1.84	25.5	31.5
	2.45	19.5	29.5
	3.06	15.5	27.0
Mean		20.2	29.3

Note. Transition times are given in msec; V is in mm; D = 17 mm; H = 6.8 mm. Display was viewed from 2 m (see Figure 1, Panel a).

informal experiment in which the intensity of the rings was reduced in several steps while transition points were remeasured for subject PB. With a small intensity reduction, transition values were larger and similar to those obtained with the dot-dot stimulus. However, when stimulus intensity was reduced further, transition values fell well below those obtained in Experiment 2. It should be emphasized that if motion is preferred between like elements, energy differences and pattern differences each should decrease transition values in Experiment 3. The fact that an energy change could be introduced to cancel the effect of a supposed pattern difference suggests that it is an effective energy difference rather than pattern difference that accounts for the original result.

Discussion

We have demonstrated a phenomenon of path selection in apparent motion. When a stimulus provides several paths along which individual points may appear to move, observers generally see motion clearly over only one of these paths (Experiment 1); apparent

motion over other paths is much reduced or completely suppressed. Path dominance was found to depend in a systematic way on the time and space intervals between stimulus elements along candidate paths (Experiment 2) but very little on the figural similarity between consecutive elements along a path (Experiment 3).

We now derive a quantitative description of the time and space trade-off that determines path selection. In this analysis, we assume that path dominance is determined by an underlying quantity, which we call the *stimulus strength for apparent motion*, or *strength*. Stimulus strength is inferred from the transition data of Experiment 2.

### Derivation of Stimulus Strength

There is a potential path for motion between any two stimulus elements that are not presented at the same moment in time. However, in our experiments, apparent motion was observed only along those paths directed between a given element and the nearest element in the row presented  $i$  time intervals later (see Figure 1, Panel a). We designate these paths as  $P_i$ .

Let  $d_i$  and  $t_i$  be the distance and time intervals between successive elements on path  $P_i$ . Then  $t_i = it$ . With each path  $P_i$  we associate a real-valued quantity  $S_i$ , the stimulus strength for apparent motion along path  $P_i$ .

The theory rests on five assumptions. The first three of these characterize  $S_i$ , the stimulus strength for motion along path  $P_i$ , and relate it to perceived motion.

*Assumption 1: The independence of alternative paths.* Strength  $S_i$  depends only on the time and distance intervals,  $t_i$  and  $d_i$ :  $S_i = S(d_i, t_i)$ . Thus the stimulus strength for path  $P_i$  is not altered by the proximity of other paths or by factors such as retinal position (within the range of retinal positions studied).

*Assumption 2: Continuity.*  $S$  is a continuous function of  $d$  and  $t$  and a strictly monotonic function of  $d$ .

*Assumption 3: Path selection (linking hypothesis).* For any given stimulus, let  $(d_i, t_i)$ ,  $(d_j, t_j)$ ,  $(d_k, t_k)$  represent paths with path numbers  $i, j, k$ . If for all  $j \neq i$ ,  $S(d_i, t_i) > S(d_j, t_j)$ ,

then the dominant motion (when motion is seen) is perceived to be on path  $P_i$ . It follows that at the transition point between paths  $P_i$  and  $P_j$ ,  $S(d_i, t_i) = S(d_j, t_j)$  and  $S(d_i, t_i) > S(d_k, t_k)$  for all  $k \neq i, j$ .

This assumption links the theoretical construct of stimulus strength to the perceptual response that was measured. Two auxiliary assumptions describe the transition data of Experiment 2.

*Assumption 4: Scale invariance.* Transition times do not depend on stimulus scale; that is, they do not vary with viewing distance. It follows that transition times  $t_{i,j}$  are a function of the distance ratio  $d_j/d_i$ .

*Assumption 5: Log linear dependence.* The transition times  $t_{i,j}$  are a linear function of the log of the distance ratio  $d_j/d_i$ . In particular, given that Path  $i$  is of order 1 and Path  $j$  is of order  $n$ , there are positive variables  $A(n)$  and  $B(n)$  such that

$$t_{1,n} = A(n) - B(n) \log(d_n/d_1). \quad (1)$$

The overall plan is to develop a theory that accounts for the main features of the data and then to modify the theory to account for second-order features. Thus, we have already discussed scale invariance and have noted that the data exhibit a small but significant deviation from scale invariance. However, we adopt scale invariance as a simplifying assumption for this initial derivation of the theory. Later, the assumption will be relaxed, and we will modify the theory to account for the deviation from perfect scale invariance.

The log linear relation expressed in Assumption 5 is a good approximation to the data of Experiment 2 (Figure 3). It is not the only possible description. The assumption is adopted because it gives an adequate description of the data and also leads to a simple form for  $S$ .

*Scale invariance implies  $S(d, t)$  is a separable function of  $d$  and  $t$ .* In the context of our experiments, scale invariance, Assumption 4, means the following. Assume that two paths  $P_1$  and  $P_2$ , with time and distance parameters  $(d_1, t_1)$ ,  $(d_2, t_2)$ , are at their transition point; that is, they have equal strength (Assumption 3). Then for any magnification or minification  $\alpha$ ,  $\alpha > 0$ , of the stimulus,

$$S(d_1, t_1) = S(d_2, t_2) \Leftrightarrow S(\alpha d_1, t_1) = S(\alpha d_2, t_2). \quad (2)$$

That is, the strength equality of  $P_1$  and  $P_2$  at their transition point, which is described by the left side of Equation 2, implies, and is implied by, the strength equality at the transition point described by the right side of Equation 2.

In Appendix A, Equation 2 is shown to belong to a class of translation functional equations whose solutions (Aczel, 1966, chap. 6; Falmagne, Note 2) are of the form

$$S(d, t) = H(d^{-1}g(t)), \quad (3)$$

where  $H$  is any strictly increasing monotonic function. That  $d$  occurs raised to a negative power in Equation 3 is a consequence of the following finding in our data: When apparent motion along a particular path is dominant, then increasing the  $d$  associated with that path may cause the motion to become suppressed. However, the reverse does not occur; decreasing  $d$  never produces suppression. Therefore,  $S(d, t)$  must be a decreasing function of  $d$ . In fact, this property of  $d$  was not tested directly but follows immediately from the data.<sup>2</sup>

Equation 3 can be understood as follows. Strength is known only in an ordinal sense. That is, if  $S(d, t)$  satisfies Assumptions 1–5, then  $S' = H(S)$  will also satisfy Assumptions 1–5, where  $H$  is any strictly increasing monotonic function. This indeterminacy in  $S$  is a consequence of the fact that the data of Experiments 2 and 3 represent only balance conditions in which two path strengths are equal. From balance conditions one derives only an ordinal scale, and this is explicitly expressed in Equation 3. Further, if  $g(t)/d$  satisfies Equation 3, then  $g^m(t)/d^n$  will also satisfy Equation 3 for  $m, n > 0$  because the constants  $m, n$  can be absorbed into the definitions of the functions  $g$  and  $H$ .

Since the units we use to measure  $S$  are immaterial for our present purposes, we absorb the monotonic function  $H$  into  $S$ , and we have

$$S(d, t) = g(t)/d, \quad (4)$$

where  $S$  is now defined only up to an arbitrary monotonic transformation.

Equation 4 represents an extremely pow-

erful result. That strength can be expressed as the product of a function  $f$  of distance and a function  $g$  of time [ $S(d, t) = f(d)g(t)$ ] is called *separability*. It means there is no interaction between the time and distance contributions to strength. For example, it means that when an optimum  $t$  (for strength of apparent motion) is found for one interdot spacing  $d$ , it will be the optimum  $t$  for every interdot distance. Similarly, if an interdot distance  $d_1$  produces greater strength than  $d_2$  at one value of time  $t_1$ ,  $d_1$  will produce greater strength at every value of  $t$ . These relations are summarized in Equations 5a and 5b.

$$S(d_1, t_1) > S(d_1, t_2) \Leftrightarrow S(d, t_1) > S(d, t_2) \text{ for all } d, \quad (5a)$$

$$S(d_1, t_1) > S(d_2, t_1) \Leftrightarrow S(d_1, t) > S(d_2, t) \text{ for all } t. \quad (5b)$$

It is trivial to verify that a separable function of the form  $g(t)/d$  will satisfy scale invariance and the remaining assumptions. The remarkable demonstration of this section is that Assumption 4 (together with the weak restrictions of Assumptions 1–3) implies that the time–distance trade-off in apparent motion is of this separable form.

*The time function,  $g(t)$ .* Having discovered the form of the time–distance trade-off in  $S(d, t)$  we still need to determine the time

<sup>2</sup> To show that  $S(d, t)$  is a decreasing function of  $d$ , consider a graph of transition points  $t_{1,n}$  versus  $d_n/d_1$ , such as Figure 3. Whenever the locus of  $t_{1,n}$  is a curve (approximated as a straight line in Figure 3) with negative slope, points above and to the right of the locus will represent  $P_1$  dominance, and points below and to the left will represent  $P_n$  dominance. In fact, all our  $t_{1,n}$  versus  $d_n/d_1$  loci have a negative slope.

Now consider a horizontal line ( $t = c$ ) that intersects the locus. Suppose it were possible to do an experiment in which  $d_1$  varied as  $d_n$  (and  $t$ ) remained fixed, that is, to have stimulus parameters move along the horizontal line to the right in Fig. 3. If  $d_1$  were decreased, the region of  $P_1$  dominance would be reached. Thus smaller  $d_1$  favors  $P_1$  dominance. Similarly, in an experiment in which only  $d_n$  varied, decreasing  $d_n$  (leftward movement along the horizontal line) favors path  $P_n$ . Thus decreasing each of  $d_1$  and  $d_n$  results in increased dominance of the associated apparent motion. By the linking Assumption 3, stimulus strength must be increasing as  $d$  is decreased to produce the increased dominance. Therefore the exponent of  $d$  in Equation 3 must be negative.

function,  $g(t)$ . The derivation of  $g(t)$  requires Assumption 5, which has not been used until now.

According to Assumption 3, at a  $P_1$ - $P_n$  transition point, there is strength equality; that is,  $S(d_1, t_1) = S(d_n, t_n)$ . We note that  $t_1 = t_{1,n}$ ,  $t_n = nt_1$ , and that according to Equation 4 we can substitute  $g(t)/d$  for  $S(d, t)$ . This yields

$$g(nt)/g(t) = d_n/d_1.$$

Let  $h(t) = \log [g(t)]$ . Then

$$h(nt) - h(t) = \log (d_n/d_1).$$

The right-hand side of this equation is now in a form similar to the rightmost term of Equation 1 (Assumption 5); substituting it into Equation 1 and rearranging terms yields

$$h(nt) - h(t) = [A(n) - t]/B(n). \quad (6)$$

Equation 6 can be reduced to a generic form of the Cauchy functional equation (Aczel, 1966, p. 41). The details of the mathematical derivation are outlined in Appendix B; the solution for  $h(t)$  is

$$h(t) = \alpha\beta \log t + \beta(1 - t) \quad (7a)$$

$$A(n) = (\alpha \log n)/(n - 1) \quad (7b)$$

$$B(n) = [\beta(n - 1)]^{-1}. \quad (7c)$$

Here,  $\alpha$  and  $\beta$  are positive constants independent of  $n$  and  $n \geq 2$ . Equations 7b and 7c mean that Equation 6 implicitly places powerful constraints not only on  $h(t)$  but also on the empirical parameters  $A(n)$ ,  $B(n)$ . When Equation 6 is true, (i.e., when Assumptions 1 to 5 adequately describe the data), then  $A(n)$  and  $B(n)$  must vary with  $n$  as described by Equations 7b and 7c.

Now the time component function  $g(t)$  can be recovered:

$$g(t) = \exp[h(t)] = t^{\alpha\beta} e^{-\beta t} e^{\beta}. \quad (8)$$

The function  $g(t)$  is nonnegative, has a single relative maximum at  $t_{op}$

$$t_{op} = \alpha = (n - 1)A(n)/\log(n), \quad (9)$$

and  $g(t)$  is zero at  $t = 0$  and in the limit as  $t \rightarrow \infty$ . Equation 9 demonstrates that in principle, the time  $t_{op}$  that optimizes strength of apparent motion can be estimated from a single transition point obtained when  $d_1 = d_n$ , since any such point suffices to de-

fine  $A(n)$ , the right edge intercept, Figure 3. Or, the optimum time  $t_{op}$  can be estimated from any two  $P_1$  to  $P_n$  transition points (provided  $d_n/d_1$  is not the same for both), since any two such points define a line whose intercept  $A(n)$  can be used in Equation 9. In practice, of course, it is better to obtain such estimates from more data, as we do below.

*Prediction of  $P_1$  to  $P_n$  transitions: Basic theory.* Combining Equations 4 and 8 gives the stimulus strength function that constitutes the basic theory:

$$S(d, t) = t^{\alpha\beta} e^{-\beta t} / d, \quad \alpha, \beta > 0. \quad (10)$$

Because  $S(d, t)$  is defined only up to a monotonic transformation, all multiplicative constants (such as  $e^\beta$ , Equation 8) have been absorbed into  $S(d, t)$ . Equation 10 describes the time-distance trade-off strength of apparent motion. It is precisely this trade-off that determines transition times of Experiment 2.

All the information needed to predict transition times is already contained in Equations 7b and 7c. Substituting these values for  $A(n)$ ,  $B(n)$  into Equation 1 yields

$$\frac{t_{1,n} = \alpha \log(n) - \beta^{-1} \log(d_n/d_1)}{n - 1}, \quad (11)$$

where  $d_n$  represents interdot distance along a path of order  $n$ , and  $d_1$  represents a path of order 1.

Equation 11 has only two free parameters  $\alpha$ ,  $\beta$  but it may be used to predict all 36 observed transition points. We found values of  $\alpha$ ,  $\beta$  that yield a least squared error fit of this theory to the data. The resulting predicted transition times are shown as the straight lines in Figure 3. The values of  $t$  for which  $S(d, t)$  is maximum,  $t_{op}$  ( $=\alpha$ ), are 18 and 24 msec for subjects PB and GI, respectively.

There is qualitative agreement between theory and data in two respects: in the rates at which transition times decrease as the distance ratio is increased and in the relative values of transition times for the three transition types. Nevertheless, there is a clear and systematic error between theory and data. The largest part of this discrepancy occurs for data obtained when one or both paths have very small  $d$ , that is, for  $d$  less than 6 minutes of arc. The discrepancy may



be traced to an unrealistic feature of the stimulus strength expression of Equation 10, namely that  $S(d,t)$  increases to infinity as  $d$  approaches zero. Clearly, strength cannot continue to increase once  $d$  is less than some spatial resolution limit of the system.

*Modified theory.* To incorporate the limit of spatial resolution in  $S$ , we multiply the distance function by a correction factor,  $\exp(-\gamma/d)$ , which has little effect at large  $d$  but which forces  $S$  to 0 at  $d = 0$ . In fact, we tried several different correction factors and found that the predictions were not sensitive to the form of the correction factor. What matters is that the corrected  $S(d,t)$  approaches a reasonable limit, not infinity, as  $d$  goes to zero. The factor chosen,  $\exp(-\gamma/d)$ , is typical of those that work well but it has no particular theoretical significance. The modified strength  $\hat{S}$  is given by

$$\hat{S}(d,t) = t^{\alpha\beta} e^{-\beta t - \gamma/d} / d. \quad (12)$$

Figure 6 illustrates the effect of this modi-

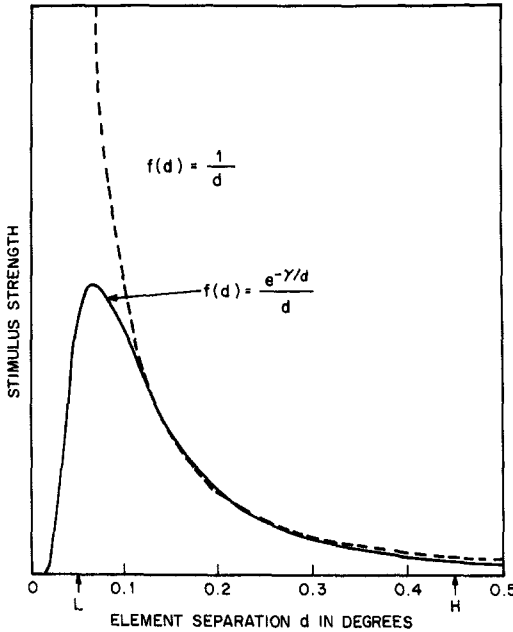


Figure 6. The dependence of the stimulus strength for apparent motion on element separation. (The distance contribution to stimulus strength is shown on the ordinate for the initial model [Equation 10, dashed curve] and the modified model [Equation 12, solid curve] as a function of element separation  $d$ , on the abscissa. Curves are normalized to cross at  $.14^\circ$ . The low [L] and high [H] limits to the range of  $d$  represented in the data are indicated by arrows.)

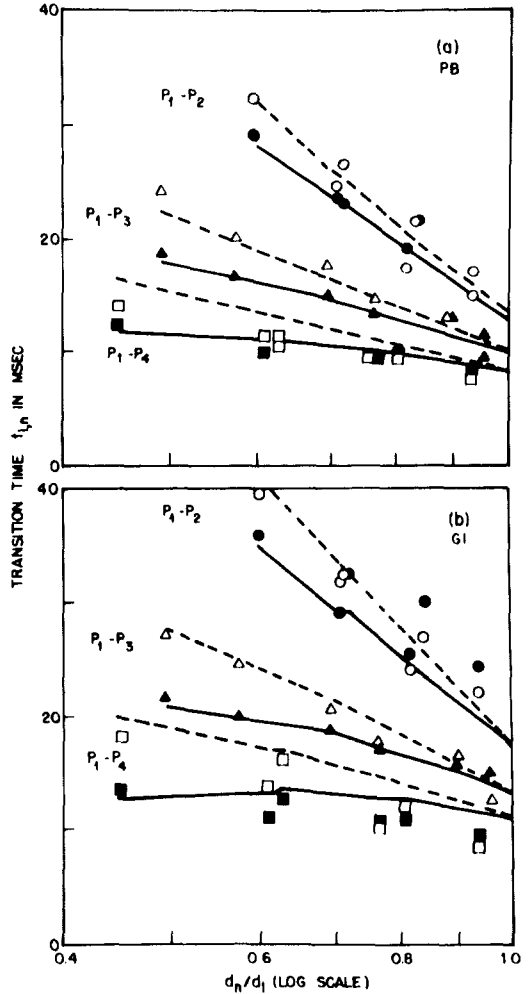


Figure 7. Predicted transition times for the modified model. (Experiment 2 data are replotted as in Figure 3. The predicted transition times based on Equation 12 are joined by solid curves for the 2-m viewing distance and dashed curves for the 1-m distance.)

fication to the  $S$  function. The modified theory no longer exhibits perfect scale invariance, but this modification results in a substantially improved agreement between prediction and observation, as shown in Figure 7.

The modified theory accounts for 96% and 93% of the variance in the transition-point data for subjects PB and GI, respectively. The average absolute values of the prediction errors for the two subjects are 1.3 and 2.2 msec, respectively. These values are comparable to the estimated measurement errors. Nevertheless, small but systematic de-

viations from data still occur, particularly for the  $P_1$  to  $P_2$  transitions at large  $d_2/d_1$ , and for  $P_1$  to  $P_4$  transitions at small  $d_4/d_1$  ratios.

*Parameters.* Equation 10 has two parameters  $\alpha$  and  $\beta$ . These suffice to describe the time-distance trade-off in the strength function and thereby to predict the transition points. Since strength is determined only up to an arbitrary monotonic transformation by our data, the parameters clearly do not suffice to determine strength itself, only the time-distance trade-off in strength. For example, raising the right-hand side of Equation 10 to any power would still preserve it as a candidate strength function; it would not alter any of the transition-point predictions, nor  $t_{op}$ , nor any of the trade-off properties under consideration.

The modified theory contains an additional parameter  $\gamma$ , which was optimized in the predictions. The parameter  $\gamma$  primarily influences predictions for small values of  $d$ , and it determines where the relative maximum of the distance function  $f(d)$  occurs. In fact, our data do not enable us to say much about  $f(d)$  for small  $d$ s except that, insofar as  $f(d)$  has a maximum, it occurs for  $d$  approximately at or below the smallest distances studied.

*Review of the theory and its relation to previous work.* The modified form of the stimulus strength function, Equation 12, captures the observed motion selection behavior over the range of  $d$  and  $t$  for which we have obtained data. At the extremes of this range, and outside, apparent motion is of poor quality even for unambiguous stimuli. Thus the good match of theory to data within the range of  $d$  and  $t$  under study is critical, whereas possible mismatches outside this range would be unimportant.

The critical properties of the theory are:

1. *Separability.* Time and distance make independent contributions to stimulus strength,  $\hat{S}$ .

2. *Scale invariance.* A change in the overall scale of the stimulus does not affect path selection, provided distances are larger than about 6 minutes of arc.

3. *Optimal  $t$  and  $d$ .* There is an inter-stimulus interval,  $t_{op}$ , which maximizes the stimulus strength for any path, irrespective of the spatial separation of dots along the

path. This optimal  $t$  is approximately 20 msec. Similarly, there is an optimal  $d$  of about 6 minutes of arc or less.

A fourth property follows from Experiment 3:

4. *Feature insensitivity.* The stimulus strength is not affected by figural details of the stimulus elements; there is no consistent tendency to see apparent motion between like rather than unlike elements.

Our experiment and analysis have led to an idealized, quantitative statement of the contributions of time and distance to apparent motion. We know of no previously published results that are comparable. The separability of space and time and the discovery of an optimal  $t$  are in direct contrast to the conclusion drawn by Korte (1915) that the optimal time interval is proportional to the separation of components in the two-component stimulus. Kolers (1972), in his discussion of data published by Neuhaus (1930), observes a relationship between time and distance in apparent motion that is qualitative but otherwise in agreement with our results. Neuhaus measured the range of  $t$  that produced optimal apparent motion. If we take the midpoint of this range as an estimate of the optimal  $t$ , then this value is found to be roughly independent of stimulus separation. Interestingly, the  $t_{op}$  computed in this way is about 200 msec—an order of magnitude larger than the value we obtain! This large difference between Kolers's conclusion and our own can be attributed to the differences in stimuli used in these studies: Neuhaus used a classical two-flash stimulus, whereas we used a regular sequence of many flashes.

We know of only one previous parametric study of apparent motion in many-flash presentations—that reported by Sperling (1976). Sperling's observers judged the quality of apparent motion of a spot in two-flash and in many-flash presentations. It is difficult to compare those results to ours because in Sperling's (1976) procedure, all paths with different  $d$ s and  $t$ s are precisely superimposed, and the quality of motion represents the sum of many paths rather than the outcome of a competition between paths as in the present experiments. For high-quality motion (many points, small  $d$ ,  $t$ ), Sperling found optimal quality at an interspot interval

of about 30 msec, quite similar to the present results. For low-quality motion (few points, large  $d$ ,  $t$ ), the  $t_{op}$  increased, and Sperling's two-spot data were consistent with those of Neuhaus (1930) and others who found a  $t_{op}$  of 200 msec and greater. The present theory does not predict any changes in optimum  $t$ ; we do not know whether these different outcomes of two-view and many-view experiments are due to involvement of a different class of motion detectors or to some other process.

The time-distance separability we find for apparent motion is evident also in contrast detection thresholds for a moving sine wave grating as reported by Koenderink, Bouman, Bueno de Mesquita, and Slappendel (1978). These investigators found little change in the optimal temporal frequency with eccentricity, although the optimal spatial frequency changed systematically. Time-distance separability is indicated by the fact that the spatial characteristics change independently of temporal characteristics (which do not change). Time-distance separability is also suggested by results from nonmoving but temporally modulated stimuli. Wilson (Note 3) finds psychophysical evidence that the spatiotemporal response of a transient mechanism in humans may be expressed as a product of separate spatial and temporal functions. Similarly Tolhurst and Movshon (1975) report that simple and complex cells in cat striate cortex show a separable response to temporally modulated sine wave gratings.

*Feature insensitivity.* Our observation that figural details of elements have little effect on apparent motion is consistent with previous reports. Kolers (1972) found little preference for apparent motion between like elements in a two-flash stimulus that contained potential paths for motion between like and unlike elements. Navon (1976) substantially elaborated Kolers's procedures and found that "figural identity does not have any effect in determining the type of motion experienced except when the figural analysis involved is very simple" (p. 130). Curiously, Navon reported that brightness did not affect the path of apparent motion—a finding that contradicts our observation. A possible resolution is offered by Pantle and Picciano's (1976) finding that apparent

movement from a light to a dark spot differed from apparent movement between two light spots when the interspot time  $t$  was less than about 50 msec, but not when it was longer. Thus, brightness matters under the conditions of our experiment (short  $t$ ) but not under the conditions of Navon's (long  $t$ ). Anstis (1970) found that when motion is evoked by random texture patterns it is determined on a local basis by brightness changes rather than on a more global basis by element shape. The common conclusion of these quite diverse experiments is that featural similarity has a small effect on apparent motion relative to interelement time and distance. Insofar as featural differences have had an influence on motion perception, they seem to have been due to very gross features, such as brightness or, perhaps, size. At long  $t$ , there is not even an effect of brightness.

### *Neural Mechanisms*

The question we propose to answer in this section is, What do the critical properties outlined above imply about the organization of mechanisms serving motion perception? We postulate that the initial stage of motion analysis—motion detection—is performed by motion-detecting units. The properties of the stimulus strength function derived above (scale invariance, time-distance separability, optimal  $t$ , and feature insensitivity) reflect the properties of these units.<sup>3</sup> A second stage of analysis is postulated to deal with the problem of stimulus matching, that is, with the inherent ambiguity of the output of the first stage. At the second stage, motion-detecting units interact, for example, by mutual inhibition. Path selection is a consequence of this interaction.

*Motion-detecting units.* The basic molecular component of our model is a motion-

<sup>3</sup> In another class of motion perception models, a wave of neural activity is set up by the moving stimulus within a homogeneous, retinotopically organized, neural network. A common feature of these wave models is the assumption that there is an inherent propagation velocity for the wave of neural activity through the neural medium (see, for example, Seelen, 1973). This wave velocity is manifest perceptually as an optimal velocity for observing apparent motion. As we obtained evidence not for an optimal velocity but rather for an optimal time interval between stimuli regardless of velocity, our results do not support this type of model.

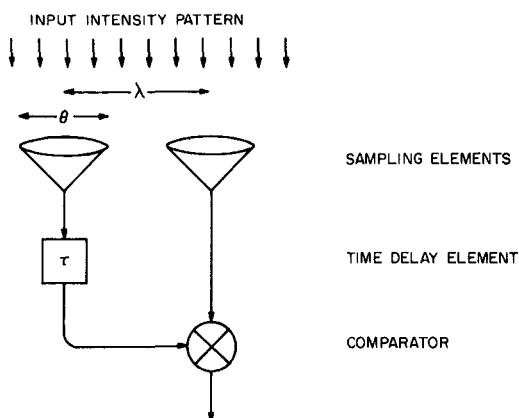


Figure 8. Motion-detecting unit. (This basic unit consists of three kinds of subunits: two image-sampling elements, a time delay element, and a comparator element. The separation between sampling element receptive fields is  $\lambda$ , whereas the size of the receptive fields, and hence the maximum size of pattern sampled, is  $\theta$ . The time delay element produces an average delay of  $\tau$  msec in signals from the left sampling element.)

detecting unit of the kind illustrated schematically in Figure 8. We propose that there are a great many such units scattered throughout the visual field, each of which responds to stimuli within its own small receptive field.

The basic motion-detecting unit consists of two image-sampling elements, a time-delay element, and a comparator element. The sampling elements send a signal to the comparator element whenever the appropriate image falls on their receptive fields. Signals from one of the sampling elements (e.g., the one on the left in Figure 8) are delayed by a time  $\tau$  ( $\tau$  represents the average delay of a filter, not a pure delay). The comparator element compares the signals from its two inputs and produces—or increases—its output when its inputs happen to match. In the example of Figure 8, this would occur when an adequate stimulus pattern moved from left to right over the sampling elements at a velocity  $v = \lambda/\tau$ , where  $\lambda$  represents the distance between the sampling units. The basic scheme of the motion-detecting unit is similar to that proposed by Reichardt (1969) for insects. Although this is not the only possible type of basic unit, we believe it is a plausible type.<sup>4</sup>

We should stress that our motion-detecting unit is proposed as a minimal motion analysis unit. There are many possible elab-

orations that do not change the basic principles of operation. For example, a comparator element could have inputs from more than two image-sampling elements; each image-sampling element could send signals to a number of correlation elements; and so forth. Our discussion of the minimal motion unit also applies to such elaborations.

*Stimulus strength.* We postulate that different populations of motion-detecting units respond to motions at different velocities and in different directions. Within a population tuned to the same direction, individual units may differ from one another in three respects: in the pattern sensitivity and receptive field area of their sampling elements, in the separation  $\lambda$  of these receptive fields, and in the time delay  $\tau$  of their delay elements. The integrated output of the entire population of units that respond to a particular moving image corresponds to the stimulus strength function we have obtained from empirical observations.

Certain important properties of motion units can now be inferred from our observations. We indicate these properties only in qualitative terms. (A complete specification of the model, which would apply to the perception of complex spatial patterns and to real motion, as well as to dots in apparent motion, is beyond the scope of the present analysis.)

*Scale invariance.* We have shown scale invariance in path selection with stimuli composed of dots. In other experiments that we have not formally reported, we found scale invariance with stimuli composed of other types of micropatterns, such as those used in Experiment 3. In order for scale invariance to hold for all stimuli, it must follow that units which respond optimally to a particular stimulus when its scale is changed are simply scaled replicas of units that responded optimally to the original stimulus. (In a scaled replica, the distance  $\lambda$  between receptive elements and the scale of the optimal stimulus pattern differs by a constant

<sup>4</sup> An alternative type of motion unit requires a single image-sampling element that responds to the local gradient in image intensity. The unit interprets a temporal change in intensity as motion in the direction of the gradient. This principle was proposed by Limb and Murphy (1975) for computer analysis of scenes containing moving objects.

factor from the original unit.) The argument behind the assertion of scale replication for detection units is that if different kinds of units were to respond to stimuli at different scales, these unit differences would ultimately reveal themselves as a preference for different paths at the different scales—for at least some stimuli. Thus the principle of scale invariance in path selection ultimately implies a corresponding principle of scale replication in motion-detecting units.

There is a structural economy in such a scale-replicated system: Units sensitive to a fixed set of image pattern types are simply replicated over a large range of scales. An additional economy is that precisely the same kinds of motion detectors serve central and peripheral vision; only the relative distribution of the numbers of different-sized detectors need change to subserve the different requirements. For example, one might expect to find fewer small-scale units in the periphery, insofar as peripheral vision is less concerned with slow movements or with small objects.

*Separability.* Time–distance separability that we observe in our empirical results implies that the spatial parameters of the hypothesized motion detectors are not correlated with their temporal parameters. In particular, if we group units according to their pattern sensitivity and field separation  $\lambda$ , then all such populations must contain the same distribution of delay times  $\tau$ . This fact suggests a further intriguing possibility: All detectors could have exactly the same temporal characteristic and differ only in spatial characteristics.

*Sampling time, real motion.* Individual motion units sample the stimulus at more or less discrete intervals,  $\tau$ . We equate the value of  $\tau$  that best characterizes the populations of units with  $t_{op}$  ( $\approx 20$  msec), the stimulus presentation interval that maximizes stimulus strength. This 20-msec sampling time applies as well when the stimulus contains a continuously moving image (real motion) as when the image moves in discrete steps (stroboscopic motion). The many units responding to continuous image motion will not be in phase, so the cumulative response will be effectively continuous.

An important implication of this characteristic sampling time is that units that have

a large  $\lambda$  and are sensitive to large-scale patterns will respond to large velocities; units that have a small  $\lambda$  and are sensitive to small-scale patterns will only respond to small velocities. This coupling between velocity and image scale has been observed in psychophysical experiments by Breitmeyer (1973).

*Feature sensitivity.* Motion-detecting units may differ in the image patterns to which they are sensitive. In our third experiment, we found no consistent preference for motion between like stimulus elements over motion between unlike elements, such as line segments that differ in orientation. This observation admits two interpretations within the model: (a) sampling elements are insensitive to pattern detail and cannot distinguish pattern characteristics such as orientation, or (b) sampling elements are sensitive to detail, but units sensitive to very different patterns may feed a single comparator element. We favor the first of these interpretations. Pattern detail could be used to great advantage in resolving matching ambiguities (see below) if detecting units matched only like patterns. The fact that this information is not used suggests that it is not available.

Since motion-detecting units are indifferent to orientation, it seems plausible that their sampling elements are neurons with concentric receptive fields. These might include, for example, neurons with antagonistic center-surround receptive field organizations (Kuffler, 1953).

#### *Local Versus Global Factors in Motion Perception*

The motion detectors described above constitute the first level of a physiological model for motion perception. When we assume that stimulus strength represents the output of populations of detector units, then we may account for properties of the stimulus strength functions, such as scaling and separability, at this level of the model. At best, however, motion detectors can account only for local aspects of motion perception. Global aspects remain unexplained. A fundamental global task is the resolution of matching ambiguities. This may be accomplished at a second, match selection, level of processing.

*Matching ambiguity and match selection.* A motion unit will respond whenever

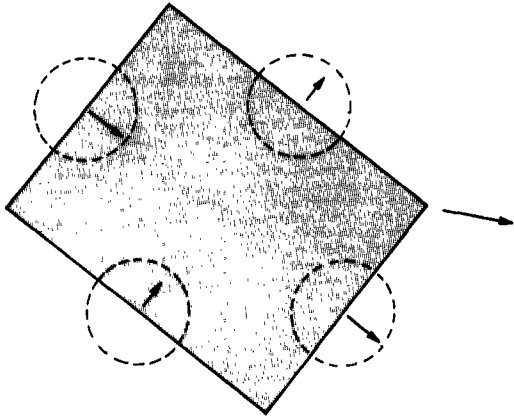


Figure 9. Demonstration of a local motion ambiguity along extended contours. (A box-shaped image moves to the right at the velocity indicated by the large arrow. This generates local motion stimuli that are perpendicular to boundary orientation, as is shown in the dashed circles. These differ from box motion and from one another in direction and velocity.)

its two sampling elements are stimulated in the correct temporal order. This may occur either when a single image feature moves across the receptive fields of both sampling elements or when one image feature crosses the first receptive field and another feature crosses the second. In the first case, sampling events are correctly matched, and the motion unit's response correctly reflects motion of an image on the retina. In the second, the unit's response results from an inappropriate match of sampling events. Since individual units cannot distinguish correct from incorrect matches they are subject to a local stimulus matching ambiguity.

In general, the output of the first motion-detecting level of processing will be a complex array of frequently contradictory motion signals. In a formal sense, the matching ambiguity can be resolved at a selection stage of motion analysis as follows: The selection stage selects a subset of unit outputs that are (a) consistent with coherent image motion and (b) complete in the sense that all sampling events are accounted for (Burt, Note 4). Details of such a system are beyond the scope of the present analysis. It should be noted that the computational task of pattern matching in motion perception is structurally equivalent to pattern-matching in stereopsis, and in that context, selection processes have been considered in some detail

(Dev, 1975, Julesz, 1971, Marr & Poggio, 1976, Nelson, 1975, Sperling, 1970).

Whatever the mechanism, selection processes should assign one and only one motion vector to each stimulus point. (The magnitude of this vector is zero for stationary images.) We suggest that the path selection phenomenon observed in Experiment 1 is a perceptual manifestation of selection processes that enforce this single motion constraint.

A particularly simple neural architecture may serve the selection we observe. In our analysis, selection was in favor of the motion path for which the stimulus strength was greatest. Suppose that populations of units that respond to motion over different paths inhibit one another at the second level of motion processing. Inhibition could suppress all detector output except that associated with the motion path with the largest stimulus strength, as required for path selection. In this interpretation,  $S(d,t)$  describes a *selection rule*, or heuristic, which the system follows in resolving the matching ambiguity.

*Other global phenomena.* In the motion system we have described, match selection is determined only by the proximity of image elements in space and time. This model is sufficient to account for perceived motion with the two-dimensional stimuli we used. It may also be sufficient for stimuli containing three-dimensional figures, where selection appears to be determined by distances in the two-dimensional image projection (Ullman, 1979).

A somewhat more elaborate mechanism must be involved in the perception of extended moving objects. This point may be made with the help of Figure 9. A tilted rectangular box is shown moving to the right. All local motion units can only detect the component of motion that is perpendicular to the contours that cross their receptive fields. Thus a single object gives rise to a complex array of local motions, which must be resolved by some further stage of analysis into coherent motion, in this case, a box moving to the right. This pattern of local motions is further complicated when several differently moving objects occur in the visual field. Not only may local motions belong to different objects, but many will be associated with features that belong to no single object

and that arise from a chance juxtaposition of objects.

Two additional levels of motion processing may interpret the array of local motion vectors. In the first, the image is segmented on the basis of the motion cue, whereas in the second, segment motion is coded hierarchically in terms of the motion relative to the background or to other segments (Johansson, 1950, Restle, 1979). These hypothesized levels of analysis are beyond the scope of the present discussion. However it should be noted that image segmentation can, in principle, be computed by a distributed, cooperative mechanism acting on the array of locally detected motion vectors. This is true because any pair of differently oriented vectors can be resolved into a unique common motion vector. Segments are formed to contain all local motions that pairwise resolve into the same common motion. (See Fennema & Thompson, 1979, and Burt, Note 4 for related discussion.)

#### Reference Notes

1. Ross, J. *Analysis before perception* (Research Report No. 4). Perth: University of Western Australia, Department of Psychology, 1972.
2. Falmagne, J. C. *Functional equations*. In J. C. Falmagne, *Foundations of psychophysical theory*. Book in preparation, 1980.
3. Wilson, H. R. *Temporal responses of mechanisms in human vision*. Paper presented at the meeting of the Association for Research in Vision and Ophthalmology, April 1978.
4. Burt, P. J. *Stimulus organizing processes in stereopsis and motion perception* (Computer and Information Science Department Technical Report No. 76-15). University of Massachusetts—Amherst, 1976.

#### References

Aczel, J. *Lectures on functional equations and their applications*. Academic Press: New York, 1966.

Anstis, S. M. Phi movement as a subtractive process. *Vision Research*, 1970, 10, 1411-1430.

Breitmeyer, B. G. A relationship between the detection of size, rate, orientation and direction in the human visual system. *Vision Research*, 1973, 13, 41-58.

Burt, P. J. Determination of the effective stimulus strength for apparent motion. *Brain Theory Newsletter*, 1976, 2, 9-12.

Burt, P. J., & Sperling, G. Contribution of spatial and temporal separation and of feature similarity to perceived movement. *Investigative Ophthalmology and Visual Science, ARVO Supplement*, 1978, 17, 289.

Dev, P. Perception of depth surfaces in random-dot ste-

reograms: A neural model. *International Journal of Man-Machine Studies*, 1975, 7, 511-528.

Fennema, C. L., & Thompson, W. B. Velocity determination in scenes containing several moving objects. *Computer Graphics and Image Processing*, 1979, 9, 301-315.

Johansson, G. *Configurations in event perception*. Stockholm, Sweden: Almqvist & Wiksell, 1950.

Julesz, B. *Foundations of cyclopean perception*. Chicago: University of Chicago Press, 1971.

Koenderink, J. J., Bouman, M. A., Bueno de Mesquita, A. E., & Slappendel, S. Perimetry of contrast detection thresholds of moving spatial sine wave patterns: II. The far peripheral visual field (eccentricity 0 degrees-50 degrees). *Journal of the Optical Society of America*, 1978, 68, 850-854.

Kolers, P. A. *Aspects of motion perception*. New York: Pergamon Press, 1972.

Korte, A. Kinematoskopische Untersuchungen. *Zeitschrift für Psychologie*, 1915, 72, 194-296.

Kuffler, S. W. Discharge patterns and functional organization of mammalian retina. *Journal of Neurophysiology*, 1953, 16, 37-68.

Limb, J. O., & Murphy, J. A. Estimating the velocity of moving images in television signals. *Computer Graphics and Image Processing*, 1975, 4, 311-327.

Marr, D., & Poggio, T. Cooperative computation of stereodisparity. *Science*, 1976, 194, 283-287.

Navon, D. Irrelevance of figural identity for resolving ambiguities in apparent motion. *Journal of Experimental Psychology: Human Perception and Performance*, 1976, 2, 130-138.

Nelson, J. I. Globality and stereoscopic fusion in binocular vision. *Journal of Theoretical Biology*, 1975, 49, 1-88.

Neuhaus, W. Experimentelle Untersuchung der Scheinbewegung. *Archiv für die gesamte Psychologie*, 1930, 75, 315-458.

Pantle, A., & Picciano, L. A multistable movement display: Evidence for two separate motion systems in human vision. *Science*, 1976, 193, 500-502.

Reichardt, W. Movement perception in insects. In W. Reichardt (Ed.), *Processing of optical data by organisms and by machines*. New York: Academic Press, 1969.

Restle, F. Coding theory of the perception of motion configurations. *Psychological Review*, 1979, 86, 1-24.

Seelen, W. V. On the interpretation of optical illusions. *Kybernetik*, 1973, 2, 111-115.

Sperling, G. Binocular Vision: A physical and a neural theory. *American Journal of Psychology*, 1970, 83, 463-534.

Sperling, G. The description and luminous calibration of cathode ray oscilloscope visual displays. *Behavioral Research Methods and Instrumentation*, 1971, 3, 148-151.

Sperling, G. Movement perception in computer driven visual displays. *Behavior Research Methods and Instrumentation*, 1976, 8, 224-230.

Tolhurst, D. J., & Movshon, J. A. Spatial and temporal contrast sensitivity of striate cortical neurons. *Nature*, 1975, 257, 674-675.

Ullman, S. Two dimensionality of the correspondence process in apparent motion. *Perception*, 1979, 7, 683-693.

(Appendixes follow)

Appendix A

Sketch of a Proof That Scale Invariance Implies Time and Distance Separability

Let  $S(d,t)$  be a real-valued function of distance and time that represents the strength of apparent motion. An apparent motion path may be represented by a time-distance pair  $(d,t)$ . Scale invariance means that two motion paths  $(d_1,t_1)$  and  $(d_2,t_2)$  that are of equal strength will remain of equal strength when both distances,  $d_1$  and  $d_2$ , are multiplied by the same factor  $\alpha$ ,  $\alpha > 0$ . Thus, we assume

$$(d,t) \mapsto S(d,t) \text{ member } \mathbf{R}_+ \quad (\text{positive real numbers}) \quad (\text{A1})$$

$$S(d_1,t_1) = S(d_2,t_2) \Leftrightarrow S(\alpha d_1,t_1) = S(\alpha d_2,t_2) \quad (\text{A2})$$

for every  $d_1, d_2, t_1, t_2$ , member  $\mathbf{R}_+$ . The left side of Equation A2 implies and is implied by the right side.

Also, to simplify the derivation, it is convenient to assume that

$$S(d,t) \text{ is a monotonic, strictly decreasing function of } d \text{ for all } t. \quad (\text{A3})$$

(As noted in the text, this assumption is supported by our data within the range of distances examined.)

1. Let  $F(\alpha,S)$  be the function that gives the new strength after an  $\alpha$  scale change. That is,

$$F[\alpha,S(d,t)] = S(\alpha d,t).$$

Note that there is a simple composition rule for  $F$ :

$$F[\alpha\beta,S] = F[\alpha,F(\beta,S)]. \quad (\text{A4})$$

2. For some path  $(d_1,t_1)$ , let  $S_1 = S(d_1,t_1) = 1$  be the unit for measuring stimulus strength. Then for any  $S$  there is a  $\beta$  such that

$$F(\beta,S_1) = F(\beta,1) = S.$$

That is,  $\beta$  is the scale reduction (or magnification) needed to achieve strength  $S$  when starting from

a path with strength equal to 1. To prove this statement, let  $(d_1,t_1)$  be the parameters of  $S_1$ . For arbitrary  $d_2,t_2$  we want to find a  $\beta$  such that  $F(\beta,S_1) = S(d_2,t_2)$ . First note that there is always a pair  $\alpha_1,\alpha_2$  for which

$$S(\alpha_1 d_1,t_1) = S(\alpha_2 d_2,t_2).$$

This is clearly true if  $F(\alpha,S_1)$  and  $F(\alpha,S_2)$  do not approach a limiting value as  $\alpha \rightarrow \infty$ , since both are assumed to be monotonically decreasing functions of  $\alpha$ . If either approaches a limit, then both must approach the same limit, as otherwise assumption A2 would be violated for large  $\alpha$ . The pair  $\alpha_1,\alpha_2$  can be found near the limiting value. Then the required  $\beta$  is obtained through the composition rule, A4:  $\beta = \alpha_1/\alpha_2$ .

3. Let  $u(\beta) = F(\beta,1) = S$ . It then follows from A3 that the inverse function of  $u$  exists:

$$\beta = u^{-1}(u(\beta)) = u^{-1}(S). \quad (\text{A5})$$

And from A4 and A5 we obtain

$$F[\alpha,S] = F[\alpha u^{-1}(S),1]. \quad (\text{A6})$$

4. By definition of  $F$ ,

$$F[d,S(1,t)] = S(d,t).$$

Now applying A6 we obtain

$$F[du^{-1}(S(1,t)),1] = S(d,t).$$

Since  $u^{-1}(S(1,t))$  is a function only of  $t$ , we may rewrite it as  $1/g(t)$ . Similarly, by defining  $S_1(x) = F(x,1)$  we rewrite  $F[d/g(t),1]$  as  $S_1[d/g(t)]$ . We now have

$$S(d,t) = S_1(d/g(t)). \quad (\text{A7})$$

Thus  $S$  is a separable function of  $d$  and  $t$  in which  $d$  and  $t$  appear as in A7. By assumption A3,  $S$  is a decreasing function of  $d$ , so  $S_1$  must also be decreasing function of its argument. Equation A7 may be rewritten in terms of an increasing function,  $H(g(t)/d)$ , as we have used in the text.

Appendix B

Derivation of the Time Function of the Time-Distance Trade-Off in Strength of Apparent Motion

*The problem.* Given scale invariance, that is, that  $S(d,t) = H[d^{-1}g(t)]$ , and given the observed transition points  $t_{1,n}$

$$t_{1,n} = A(n) + B(n) \log(d_1/d_n), \quad B(n) > 0, \quad (\text{B1})$$

what can be deduced about  $g(t)$ ?

In order to find  $g(t)$ , we first cast B1 in the form of the Cauchy function equation

$$L(xy) = L(x) + L(y). \quad (\text{B2})$$



When  $L(1) = 0$ , and under certain other constraints (Aczel, 1966), the Cauchy equation has a solution  $L(z) = A \log(z)$ .

1. At a  $1, n$  transition point,  $S(d_1, t_1) = S(d_n, t_n)$ . Substituting  $S_1$  from A7 implies  $d_1^{-1}g(t_1) = d_n^{-1}g(t_n)$ , since  $H_1$  is one-to-one. For convenience, we write  $t$  for  $t_1$  (and for  $t_{1,n}$ ) and recall that  $t_n = nt$ . Substituting into B1 we obtain

$$t = A(n) - B(n) \log g(nt) + B(n) \log g(t). \quad (B3)$$

Let  $h(t) = \log [g(t)]$ . Then from B3 we obtain

$$h(nt) - h(t) = A(n)/B(n) - t/B(n). \quad (B4)$$

2. Choose  $h(1) = 0$ , which can be done without loss of generality. Note that this choice implies  $g(1) = 1$ , which is permissible; multiplicative constants can be absorbed into the function  $S_1$ . When  $t = 1$ :

$$h(n) - h(1) = A(n)/B(n) - 1/B(n). \quad (B5)$$

Using B4 and B5 to give an expression for  $h(nt) - h(n) - [h(n) - h(1)]$  yields

$$h(nt) = h(t) + h(n) + (1 - t)/B(n). \quad (B6)$$

3. A constraint on the  $B(n)$  is derived as follows. By symmetry with B6,

$$h(tn) = h(n) + h(t) + (1 - n)/B(t). \quad (B7)$$

From B6 and B7 we get

$$B(n)/B(t) = (1 - t)/(1 - n).$$

Let  $t = 2$ , and  $\beta = 1/B(2)$ . Then

$$B(n) = [\beta(n - 1)]^{-1} \text{ for all } n < 1. \quad (B8)$$

Equation B8 gives an extremely powerful constraint on the empirical constants  $B(n)$  that must hold if the assumptions are valid. Indeed, our data give estimates of  $B(n)$  that are in agreement with B8.

4. Substituting B8 into B6 yields

$$h(nt) = h(t) + h(n) - \beta(1 - t)(1 - n). \quad (B9)$$

Let

$$H(t) = h(t) - \beta(1 - t). \quad (B10)$$

Finally, substituting B10 into B9 gives the Cauchy form

$$H(nt) = H(n) + H(t),$$

whose solution is

$$H(t) = C \log(t). \quad (B11)$$

5. The time function  $g(t)$  can now be recovered. First substitute B11 into B10, and the resultant with B8 into B4 to obtain the following condition on  $C$ :

$$C = A(n)\beta(n - 1)/\log n.$$

Since  $C$  is not a function of  $n$ , we have

$$A(n) = \alpha(\log n)(n - 1)^{-1}, \quad \alpha = C/\beta. \quad (B12)$$

6. Substituting B11 and B12 into B10 gives

$$h(t) = \alpha\beta \log t + \beta(1 - t),$$

from which we obtain

$$g(t) = \exp[h(t)] = t^{\alpha\beta} e^{-\beta t} e^{-\beta}. \quad (B13)$$

There is no arbitrary multiplicative constant in  $g(t)$  because  $g(1)$  has been fixed equal to 1.

The time function  $g(t)$  is a unimodal function,  $g(t) \geq 0$ ,  $g(0) = 0$ ,  $\lim_{t \rightarrow \infty} g(t) = 0$ , with a maximum when  $dg(t)/dt = 0$  at

$$t = \alpha = (n - 1)A(n)/\log n.$$

Received November 27, 1979  
Revision received May 13, 1980 ■