

Hierarchical Motion Organization in Random Dot Configurations

Marco Bertamini
Staffordshire University

Dennis R. Proffitt
University of Virginia

Motion organization has 2 aspects: the extraction of a (moving) frame of reference and the hierarchical organization of moving elements within the reference frame. Using a discrimination of relative motions task, the authors found large differences between different types of motion (translation, divergence, and rotation) in the degree to which each can serve as a moving frame of reference. Translation and divergence are superior to rotation. There are, however, situations in which rotation can serve as a reference frame. This is due to the presence of a second factor, structural invariants (SIs). SIs are spatial relationships persisting among the elements within a configuration such as a collinearity among points or one point coinciding with the center of rotation for another (invariant radius). The combined effect of these 2 factors—motion type and SIs—influences perceptual motion organization.

An event is the natural unit of analysis for perception. An event is an actor or object that displays a behavior against a background, such as a falling object, the changing color of a traffic light, or the locomotion of an organism in the environment. In the case of motion, the background is the event's frame of reference, a coordinate system that may, itself, be moving. Because the motion of an element is defined relative to its reference frame, the element also shares the motion of the frame, and thus motion organizations are hierarchical.

Hierarchical motion organizations are inherently ambiguous. Every motion affords an infinite number of different descriptions as a consequence of the choice of different moving coordinate systems, or frames of reference. Consider the motion of our planet. Common sense and medieval astronomy take the earth as the frame of reference for all motions: the flowing of a river; the locomotion of an animal; or the motion of the sun, moon, planets, and stars. Contemporary astronomy, on the other hand, takes the sun as a moving frame of reference. Within this reference frame, the earth is described as revolving around the sun, and the moon is described as having two kinds of motion, one around the earth and a second one, common to the earth, around the sun. This newer description of the solar system is no more true in a geometrical sense than the earth-centered one. From the point of view of providing an accurate description, any frame of reference is as good as any other. However, the use of the sun as a frame of

reference for the earth and of the earth as a frame of reference for the moon promotes a better mechanistic description of this dynamical system.

The perceptual system is adept at finding useful organizations for the motions that it encounters. When observing events, people correctly identify the motions of an object (where it is going) and motions within the object (what it is doing). The first are called *common motions* because they are common to the constellation of elements or features of the object. The second are called *relative motions*, implying that they are hierarchically organized relative to the common component.

In this article, we are concerned with two aspects of motion organization. The first is how motion type influences the establishment of a moving frame of reference. The second is how the structure of the configuration—which is not motion information per se—is used to constrain motion organization. In six experiments, we compared three different types of motion (translation, divergence, and rotation) to test whether they can support the perception of common motion. We found that translational motions in space (translation and divergence) can serve as a moving frame of reference for the perception of relative motions, whereas rotational motions cannot. The reason this distinction was not observed before has to do in part with a second aspect of event perception: Spatial invariants also constrain motion organizations. Although the importance of this factor has been noted before, there was no clear definition of what the effective spatial information was, and much empirical research has been based on configurations that confounded the two factors.

Perception of Common and Relative Motion

The problem of motion organization was identified by Wertheimer (1923/1937) and was first studied by Duncker (1929/1937) and by Rubin (1927), who used displays generated by lights on a wheel. But the best known example of hierarchical motion organization is probably the point-light walker (Johansson, 1973, 1977). In this display, lights are attached to the joints of an otherwise invisible actor. Although the information available

Marco Bertamini, Division of Psychology, Staffordshire University, Stoke-on-Trent, United Kingdom; Dennis R. Proffitt, Department of Psychology, University of Virginia.

This research was supported by National Institute of Mental Health Grant MH52640-02 and National Aeronautics and Space Administration Grant NCC 2-925.

Correspondence concerning this article should be addressed to Marco Bertamini, who is now at the Department of Psychology, University of Liverpool, Eleanor Rathbone Building, Liverpool L69 7ZA, United Kingdom, or to Dennis R. Proffitt, Department of Psychology, Gilmer Hall, University of Virginia, Charlottesville, Virginia 22903. Electronic mail may be sent to m.bertamini@liverpool.ac.uk or to drp@virginia.edu.

seems quite minimal, the identity and action is effortlessly perceived by observers as soon as the actor moves. Even though the motion of the forearm is a complex path with respect to the environment, it is perceived as a simple oscillation with the elbow serving as its frame of reference. In the example of the point-light walker, the parts of the body are rigid and connected, but they are not rigid as a whole. In this case, the object is a hierarchical system with a mechanical internal motion (Johansson, 1950, 1973). Hochberg called this aspect *articulation of motion* (Hochberg, 1986). However, in the rest of the article we exclusively use the term *hierarchical motion organization*.

Another striking example of hierarchical organization was described by Restle (1979; see also Shum & Wolford, 1983). In the top part of Figure 1, three moving dots are shown, and their velocities are described by the arrows. Two dots move vertically, whereas the third one moves in an ellipse in phase with the others. Two organizations are possible: If the configuration is seen as three independently moving dots, then they would be described as illustrated in Panel A. In this motion organization, the dot in the center is moving on an ellipse counterclockwise. However, when the configuration is seen as a single group of dots, a reorganization occurs, and the dot in the center is perceived to move on an ellipse clockwise(!), as illustrated in Panel B. The three dots now share a common vertical motion and constitute a moving system (gray area) to which all three dots belong. It is possible to say that the object is not rigid, because there are relative motions within the

configuration, but at the same time the object is not perceived to be elastic. On the contrary, just as in the case of the point-light walker, the configuration is perceived to have parts, and these parts move relative to each other the way the wheels of a car spin when the car moves. Again, the term *mechanical motion*, used by Johansson (1950, 1973), seems particularly appropriate.

In all examples of hierarchical motion organization, the distinction between common motion and relative motion can be written as a sum of two vectors: $\vec{V}_a = \vec{V}_c + \vec{V}_r$, where \vec{V}_a is a vector describing the element motion with respect to some environmental coordinate system, and where the two terms on the right are the common (c) and the relative (r) motion vectors. Johansson (1950) suggested that the common motion is chosen according to a preference for the slowest velocity. Note also that the expression *extraction of common motion* seems to imply that in a configuration the common motion is found and the relative motion is the residual component, but there is no logical or mathematical need for this to be the order in which the perceptual process operates. If one can extract relative motions first, then the common motion will be specified as residual (Proffitt, Cutting, & Stier, 1979).

Types of Common Motion

The studies reported in this article were designed to assess whether all types of motion are equally likely to be seen as a common motion. Koenderink (1986) provided a useful taxonomy of the motion fields in optic flow that has a foundation in the calculus of vector fields. There are four kinds of transformations: translation (trans), rotation in the image plane (curl), divergence-convergence (div), and deformation (def). These four motion types are differential invariants and are independent of coordinate systems. Empirical research has assessed whether the visual system actually can detect and make use of such a decomposition (e.g., Kappers, te Pas, Koenderink, & van Doorn, 1996; Orban, 1992). Lappin, Norman, and Mowafy (1991) tested human sensitivity to these transformations, using a task in which observers had to choose between two alternatives, either a coherent transformation or random motion of a set of dots. They found that detectability for any coherent transformation is good, similar to the detectability of any motion (where the alternative was stationary dots), suggesting that local motion information is interrelated. However, combining transformations lowers performance, indicating lack of independence.

Three of the motion transformations—translation, rotation, and scaling (divergence-convergence)—are good candidates to provide common motion reference frames. Translation and rotation are isometric transformations that preserve topological as well as metric relationships between the parts of a figure, whereas a scaling of the whole figure preserves every metric property up to a constant multiplicative factor.

There are, however, important differences between these motion types. For translation, the sign and magnitude of the translation can be obtained from the instantaneous velocity of any single point in the field. This is not true for other types of flow that require the identification of a location in space—namely, a center of rotation or expansion-contraction. Moreover, translation and divergence do not cycle, whereas displacements produced by rotation accumulate continuously (modulo 2π). Modern versions of the original model of a motion sensor (Reichardt, 1961) are based on bilocal

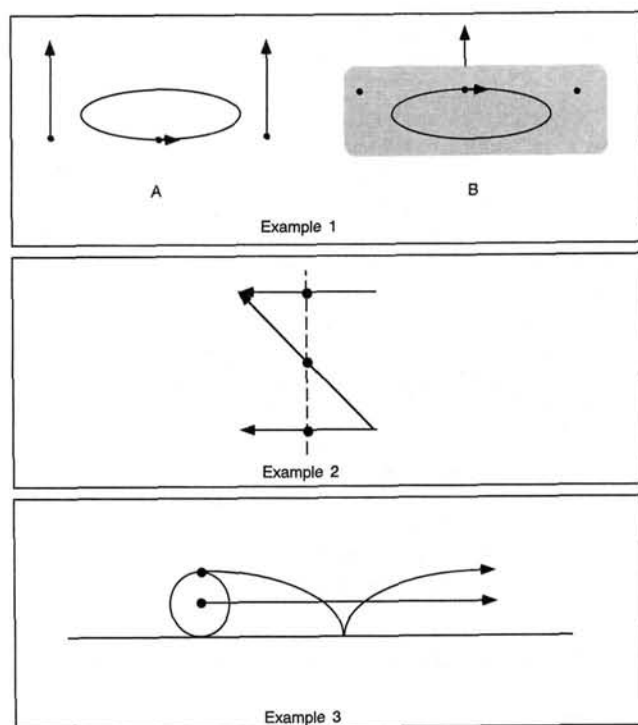


Figure 1. Example 1: Bistable hierarchical motion organization. The gray area in B represents a grouping that is perceived to move with the same velocity. The trajectory inside the gray area is a relative motion of the dot. Example 2: Three dot display demonstrating relative motion perception along a virtual line. Example 3: Trajectory of two lights on a rolling wheel, one placed on the rim and one placed on the hub.

correlations (e.g., van Santen & Sperling, 1985). Responses from such detectors extract only information about translation, and most likely such computations are carried out in the primary visual cortex (Adelson & Bergen, 1985; Kosslyn & Andersen, 1992). Only the later integration of these responses would allow the extraction of other motion transformations (Morrone, Burr, & Vaina, 1995), and this takes place in more specialized temporal areas.¹

Within the literature on motion organization, Börjesson and von Hofsten (1972, 1975) assessed the kinds of transformations that people can extract. They compared common motion (trans), circular relative motion (curl), concurrent relative motion (div), and parallel relative motion (div + def). The motion vectors were generated with a small set of dots (from two to four). They combined them and asked observers to describe the motions. They found that parallel relative motions (the only situation in which there is a deformation) led to perceived rotation in depth, and common motion (trans) led to perceived translation, even when other transformations were present. On the basis of these findings, Börjesson and von Hofsten concluded that people can extract all of these transformations.

Börjesson and Ahlström (1993) defined *common motion* as the motion of elements that do not change relative distances and *relative motion* as the change in distance between the elements, extending the classification of Börjesson and von Hofsten (1975). They used configurations of five dots in which one element can be seen as part of two different configurations, and they asked observers to report on how the elements group. They found a ranking of four types of motion based on how strongly each evoked a perceptual grouping. Parallel common motion (translation) was strongest. Then there was concurrent common motion (curl), then concurrent relative motion (div), and finally parallel relative motion (div + def). One of their conclusions, therefore, was that translation is the strongest grouping transformation in these displays with few dots, followed by rotation. Unfortunately, the composition of their stimuli, though very clever, did not allow for a comparison of basic transformations without some confounds; in particular, the rotation condition had to be compared to translation along curved paths, convergence toward a center that was not stationary, or convergence combined with slant in depth.

There is other evidence, from a completely different literature, that speaks to the topic of the difference between translation and rotation in producing grouping. Kellman and Spelke (1983) studied young infants in many clever experiments. They were interested in the comparison of different grouping principles, such as similarity, good continuation, and so on, and their development over time during the infants' first year. Common motion appears to be the most powerful grouping principle for infants who are only a few months old. This research provides evidence that parts that translate together are organized by the child as a single object, whereas parts that rotate together are not (Eizenman & Bertenthal, 1998; Kellman & Spelke, 1983).

Thus far we have considered only how different types of motion influence perceived hierarchical motions. In the following section, we discuss a second factor that relates to the spatial structure of a moving set of elements.

Structural Invariants

The way that elements are spatially organized in a configuration is important. Köhler (1947) introduced the distinction between the *dynamic* determinants of the fate of a system and its *topographical* determinants. The first have to do with the type of change over time and the second with its spatial structure. Köhler suggested that people have a preference to reason about systems in terms of mechanical devices that have topographical constraints as opposed to dynamical ones.

Duncker (1929/1937), in his seminal work, thought of *surrounding* as a critical aspect of induced motion. Surrounding is an example of structural information; namely, it is a topological property of the configuration. Pittenger and Shaw (1975) also parsed events into transformational and structural invariants (SIs). Gogel (1978) showed experimentally that the salience of any element motion depends on its perceived distance from the other elements, and he called this the *adjacency principle*. Cutting and Proffitt (1982) introduced the idea of center of moment and considered it to be an example of a structural constraint on the motion organization.

We propose a general definition: SIs consist of points or axes (defined by alignment) within the configuration that can be extracted independent of the characteristics of the motion transformation. SIs are geometrical properties that remain constant during the event; thus, they are revealed when the configuration moves. Once they become obvious, observers will be biased to detect those relative motions that are consistent with SIs.

An example of an SI is when two or more elements create a virtual line that corresponds to the motion of another element.² Consider a configuration with only three dots. Two of them are aligned vertically and translate horizontally. The third moves with the same translatory motion but also has a second component of motion orthogonal to the translation and equal in speed. Adding the two components gives an absolute motion that is at 45° as illustrated in the second example in Figure 1. The perceptual system in this case extracts the common translation so that the residual orthogonal motion of the element is perceived (vertical relative motion). However, in addition to this common motion, SI information is also present because the residual motion has a path that is along the virtual line connecting the other two elements. This correspondence is a nonaccidental property of the display.

Another example of an SI can be called *pivoting*. The importance of locating a pivot point is clear if we consider that many mechanical systems involve oscillations, or rotations around a point. The center of the rotation may be extracted from the common motion of the set of elements, but, especially in the case of few elements, these rotations are easier to organize if the pivot

¹ It is important to keep in mind that whether one transformation is simpler than another also depends on the description chosen. For instance, translations can be a subcase of rotation with a center at infinity, and in this vein, Restle (1979) used circular motions to generate a large number of possible motion displays, including linear motion.

² The fact that a virtual line may bias the organization of motion is analogous to the fact that an explicitly drawn path may bias the trajectory of apparent motion (Shepard & Zare, 1983). The analogy is based on the fact that in both cases ambiguity is resolved by relying on a nonaccidental property of the display.

point's location is visibly marked. For example, the rolling wheel display evokes a strong impression of a wheel even when only two dots are used, provided that one of the two is placed at the center of the wheel (see the third example in Figure 1).

It is important to note that SIs cannot be extracted from any single static frame of a display. Viewing a static array of point lights does not inform one about which points will maintain constant alignment or serve as pivots once the display is animated. Moreover, SIs are present even in random-dot kinematograms. SI constraints are revealed by motion; however, their status is structural, not dynamic.

General Hypotheses

The experiments assess three competing hypotheses about the influence of motion type on perceptual motion organization.

1. *Frames of reference are only perceived in configurations that translate in space:* Translation and divergence function as moving frames of reference because they specify linear displacement in the environment, whereas rotation does not.

There are at least two alternatives to be considered.

2. *A frame of reference is any type of motion transformation that can be extracted from a display:* In the event-perception literature, it is implicitly assumed that a common motion can be any kind of common transformation. In this case, there are no differences predicted. A subcase of this hypothesis is that there are only small quantitative differences. That is, although all transformations can provide a moving frame of reference, some will be more salient or will require a higher or lower speed than others.

3. *A frame of reference is more likely to be a common translation:* A completely different hypothesis is based on the mathematical and physiological differences between types of motion transformations. Motion detectors can extract translation locally because translation is local in a sense that rotation and divergence are not. Pure translations do not specify any axis or center of the field; therefore, the information about a pure translation can be recovered from any point in the field. This hypothesis predicts an advantage for translation over the other two transformations.

Because certain motion organizations may be biased by SIs, the experiments were performed with sets of randomly generated configurations where no SI confound was systematically present. Experiment 5 assessed directly our second general hypothesis that SIs are important and that it is possible to separate experimentally the effects of motion type and SIs. In all experiments, we assessed the perceived relative trajectory of a single dot moving in a closed loop (e.g., a circle) centered on an invisible point (except Experiment 5). The relative motion of this point was combined with the common motion of an array of randomly positioned dots that underwent either a global translation, rotation, or divergence.

Spontaneous Descriptions

A pilot experiment was conducted to collect and to classify spontaneous descriptions of motions in three-dot displays. The motion of one of the dots in the configuration was a combination of a pure transformation (rotation or translation) and a second component (relative motion). Consider Figure 2. These three-dot displays illustrate the phenomenon of hierarchical motion organization. The black lines are the absolute motion of the elements.

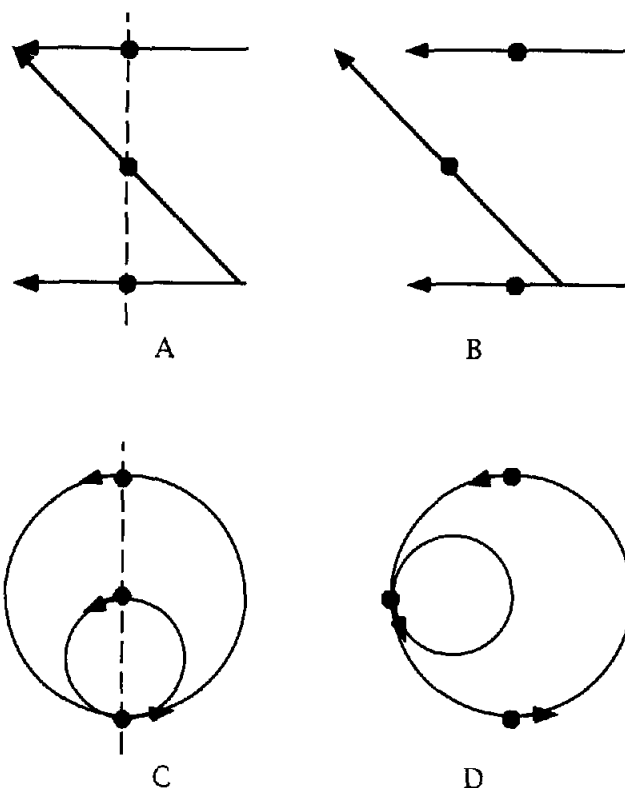


Figure 2. A three-dot display with a common translation (A and B) or a common rotation (C and D). The dashed lines show that the dots remain aligned in Conditions A and C; we call this condition SI+. SI = structural invariant.

The apparent hierarchical organization is depicted with dashed lines. In all cases, the common motion corresponds to the motion of the two outside dots. The third dot shares this motion component with the outside dots (right to left in A and B, or rotating in C and D), and in all cases the third dot also has a linear relative motion component. When the three dots stay aligned during the motion (A and C) the relative motion corresponds to a virtual line connecting the two outside dots, as described by the dashed lines. Comparing A and B with C and D allows an evaluation of the effect of alignment.

The animation was created on a Macintosh PowerPC, and the width of the display (the distance between the two outside dots) was approximately 2.5° of visual angle. The speed was 4 s per cycle, and observers could repeat the animation by pressing the space bar on the keyboard. In the left column, the three dots stay aligned during the animation, and therefore there is an SI constraint (SI+); in the right column, this constraint is not present (SI-). Each of 28 observers saw only one of the configurations (7 observers per cell) and was asked to describe the motion as accurately as possible.

The process of categorizing verbal reports is difficult and was carried out independently by two experimenters, who then met to resolve the discrepancies. Fortunately, the observed differences between conditions were large. A description of a motion of the whole configuration followed by a description of the relative

motion (between the other dots) was taken as evidence of hierarchical motion organization. A typical report of hierarchical motion organization for the translation case would be "The two end dots move left to right together and then back; the other one is going between the two." For the translation conditions, in the SI+ case, 86% of the observers organized the motion hierarchically, and in the SI- case, 43% organized the motion hierarchically. On the other hand, for the rotation conditions, in the SI+ case, 71% of the observers organized the motion hierarchically, and in the SI- case, 0% organized the motion hierarchically.

It appears that for translation the alignment can be eliminated without a qualitative change in the appearance of the display. That is, the effect is only a weakening of the hierarchical organization produced by the misalignment. A different result was found for rotation using Configurations C and D in Figure 2. Only when the motion of the middle element follows the virtual line defined by alignment could the observer see the residual linear motion. When alignment was eliminated, observers did not extract the linear component from the common rotational motion. The two factors, type of motion (advantage for translation) and SI (advantage for conditions where alignment is present), seemed to combine.

This demonstration shares some problems with other demonstrations in the literature that use few dots. First, alignment is present also as static information (i.e., a virtual line). A more serious limitation is that three dots are too few to create a non-structured configuration. The line or triangle that they form is a simple shape that may be seen as such, almost as if illusory contours were present, as opposed to a configuration in which elements are random pieces of a whole. Moreover, in the demonstration we used a single configuration per condition, and therefore it is difficult to generalize our findings.

A solution to all these limitations is to use a larger number of dots, place them at random locations within a defined area, and change this configuration to create many different displays that have in common only the particular type of motion under investigation. In the following experiments (Experiments 1-6), this kind of display was used.

Effect of Motion Type

We decided to systematically explore when a common motion transformation provides a moving frame of reference for perception of relative motions. The assumption is that when a common transformation is perceived as the motion of the whole configuration, the relative motion will be perceptually available to the observer, and therefore judgments on the trajectory of relative motion will be possible. In other words, the stronger the hierarchical organization, the easier the discrimination between two similar relative motion trajectories. To test this assumption, we performed six experiments in which observers judged the perceived shape of relative motion of a single dot while the reference frame (a set of dots) always underwent a simple transformation. When the transformation was cyclical (rotation), the animation lasted exactly one cycle.

A problem with the idea of perception of residual motion is that vector analysis, or flow decomposition, may not be complete. In other words, perception of a relative motion requires that the common motion be completely subtracted. There is evidence that suggests that only a given percentage, although high, of the com-

mon motion is subtracted (Shum & Wolford, 1983). We did not compute the percentage of absolute motion subtracted by the observers for each type of transformation and under different viewing conditions. Instead, we assumed that even when incomplete subtraction occurs (e.g., a residual circle may appear as an ellipse), the interesting question is whether the difference in perceived relative motion will be enough to allow discrimination of different residual motions. It is possible that a smaller percentage of common motion will be subtracted for some of the conditions, but this is tantamount to a weaker frame of reference effect, which is our preferred terminology.

Experiment 1 tested the difference between types of motion transformations for variable speeds. Experiment 2 tested the same difference but for a different kind of relative motion (linear as opposed to elliptical trajectories). Experiment 3 studied the effect of variable densities (number of elements present in a fixed region). Experiment 4 studied the effect of motion transformations that are combinations of simple transformations. Experiment 5 replicated the design of Experiment 2 but introduced SI constraints in the configuration of dots. Finally, Experiment 6 served as a control for a possible indirect effect of the detectability of the target dot.

General Method for Experiments 1-5

Apparatus and Stimuli

All displays were programmed on a Silicon Graphics Indigo2 Extreme workstation. The monitor has a resolution of 1280×1024 and a refresh rate of 60 Hz. The display was antialiased to obtain subpixel resolution. Interstimulus interval was kept to zero to obtain the best animation, and the speed was varied by changing the presentation time.

The dots were red on a black background and were approximately 0.072° of visual angle in diameter. The dots had the highest luminance possible in the red-green-blue (RGB) color space, whereas the background was left completely black. The dots were drawn as rings with a central round hole, one third of the diameter of the dot. This manipulation was found to be effective in reducing the probability of perception of motion of the dots on a plane different from the frontoparallel plane. The dots were positioned within a circular region centered on the monitor. The distance of the point of observation was 70 cm; therefore, this region subtends approximately 2.41° of visual angle. The display is illustrated in Figure 3, where the vectors show the velocities of the dots. The gray cross at the center of the display was not present in the experiments; it was drawn only to mark

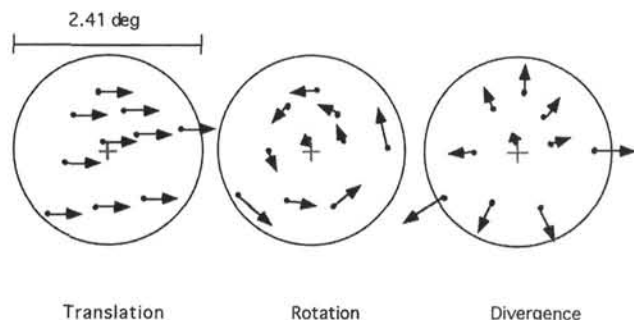


Figure 3. Common motion transformations used in Experiments 1-6. The average speed was the same for all three transformations. The range of speeds was the same for rotation and divergence. No relative motion is depicted in this figure.

the center of the circular region. Finally, the region had no visible border in the experiments.

To better describe the animation sequence, we show the trace over time of the motion of the dots in Figures 4–6. Figures 4, 5, and 6 refer to Experiments 1, 2, and 4, respectively. Within each of these figures, each display represents the trace of 10 dots on the screen and each row shows a different type of transformation. Notice that one of the dots is black and has a trajectory that is quite different from the others. The motion of this dot can be described as a combination of two motions: One component is the same as the motion of the other nine dots; the other component is an elliptical motion (rectangular in Experiment 2). This decomposition in a common and a relative motion is not the only way of describing the motion of the black dot, but—under the assumption that the black dot is seen as part of the configuration and that the configuration is taken as a moving frame of reference—the relative motion is uniquely defined by the subtraction of the common motion from the absolute motion. In the experiments, all the dots had the same color, except for Experiment 6.

The right and left columns of Figures 4–6 show the two types of relative motions that observers were asked to discriminate. Note that for rotation it is impossible to individuate the starting location of the dots, but this is only a feature of this figure, and it is not a real difference between the different motion conditions in the experiments. Several parameters in generating the displays were randomly chosen in each trial; they were (a) the location of

all the dots within a circular region of the screen of radius 100 pixels (2.41° of visual angle); (b) the location of the center of the elliptical motion; (c) the starting point along the elliptical path, which we can call *phase*; (d) the orientation of the main axis of the ellipse; and (e) the length of the axes for the ellipse (but the aspect ratio was fixed). Note that, when we talk about *elliptical motion* here and later, this applies also to the case of circles as a special case of ellipses.

Some constraints were applied to the randomization process; in particular, the center of the elliptical motion was always in the outer ring of the circular region (outside a central circle of radius 50 pixels, or approximately 1.2° of visual angle), and the major axis of the ellipse was always between 0.48° and 1.80° of visual angle. Moreover, the dots were not allowed to overlap in the configuration; some overlap may occur only because of the relative motion of one of the dots. These constraints apply to the configurations used in all the experiments, even when linear paths instead of elliptical paths were used. In that case, the values of the axes of an ellipse became the sides of a rectangle.

Procedure

During a practice session, observers were instructed to notice that one of the dots always moved differently from the others. They were asked to judge the trajectory of that dot with respect to the other dots. The judgment

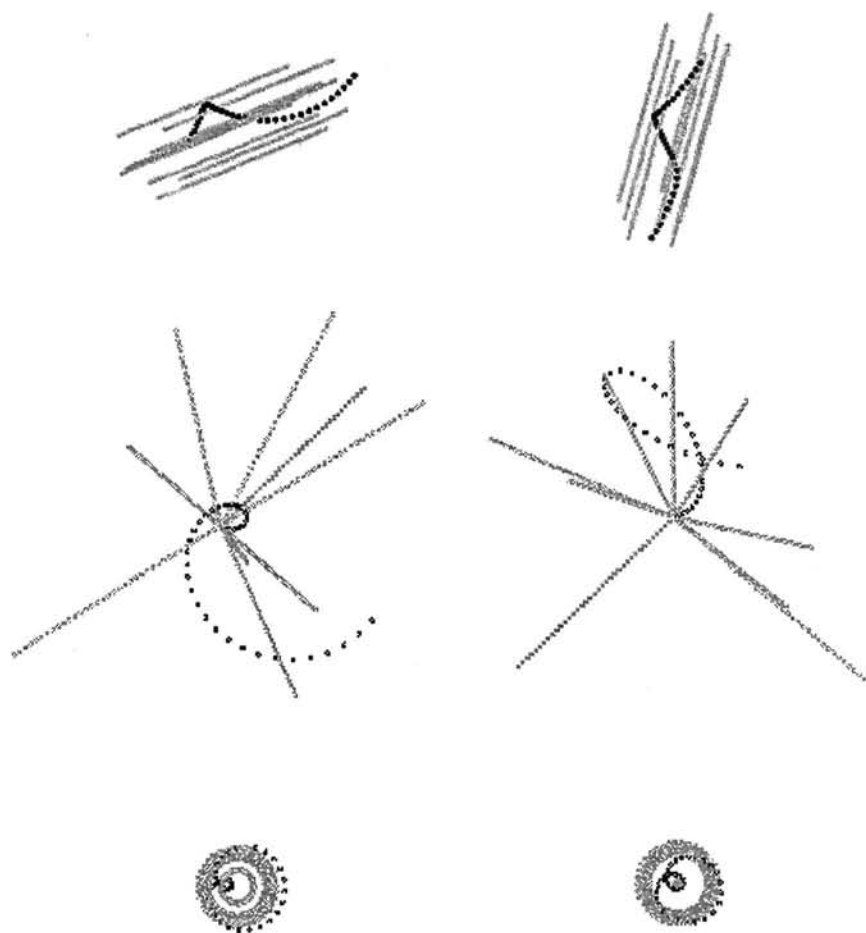


Figure 4. Trace of 10 dots during animation: examples randomly chosen from the set of all possible displays for Experiment 1. The rows have three different types of common motion: translation, divergence, and rotation. In the right column there is a (black) dot with a circular relative motion, in the left column there is a (black) dot with elliptical relative motion.

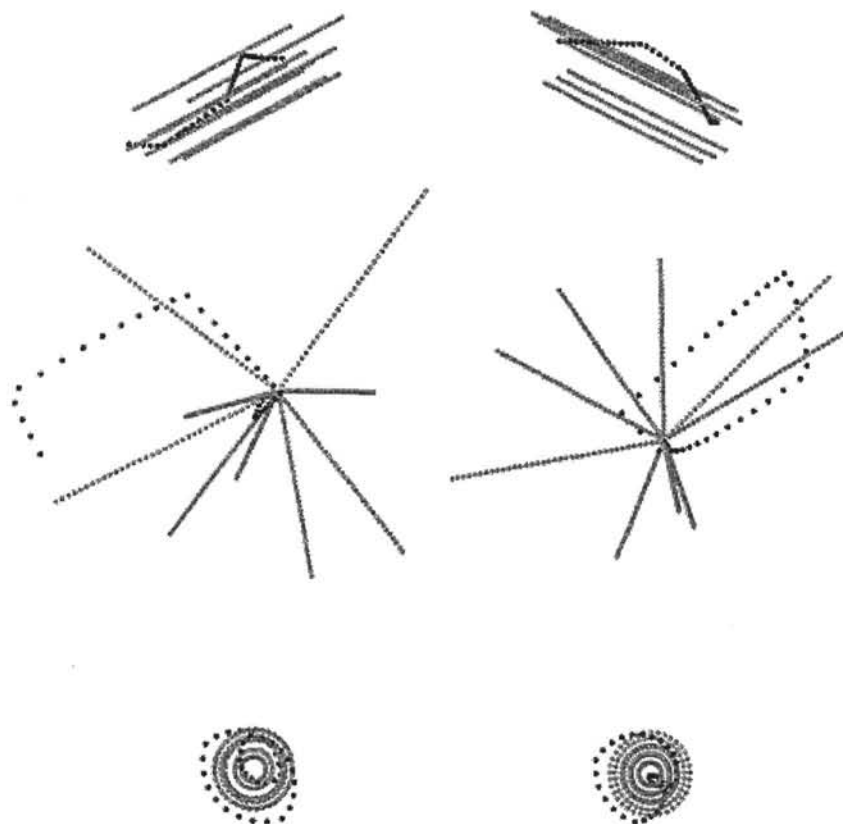


Figure 5. Trace of 10 dots during animation: examples randomly chosen from the set of displays for Experiment 2. The common motion is as in Experiment 1. The relative motions are along a square (on the right) or a rectangle (on the left).

was made by choosing one of two alternatives, presented on the screen at the end of each trial so that the observer could press one of two keys to enter the response. The screen had a single line of text, reminding the participants, for instance, that the A key meant circle and the S key meant ellipse. The keys were chosen so that an observer could use only one hand and did not need to look down at the keyboard after positioning that hand. The observers were told that there was always a correct response, but they were not given feedback during the experiment.

Observers sat in a dimly illuminated and quiet room. A small desk light was on but facing the wall so that illumination was kept at the minimum necessary to see the keyboard while avoiding any glare on the glass of the monitor. The length of the practice varied but was at least 12 trials long and always contained at least one example of each of the three motion transformations. When observers were ready to see a new trial, they pressed the space bar on the keyboard. That is, the rate of presentation of the trials was self-paced. Participants were invited to find a comfortable position in front of the monitor and to maintain it during the blocks, but no constraint was placed on their body, so that the distance from the screen was only approximately the same for all of them.

For Experiments 1–5 the observers were undergraduate or graduate students at the University of Virginia. All of them were naive with respect to the hypotheses and the design of the experiment, but the graduate students were relatively experienced psychophysical observers. Two of the observers participated in all experiments. We did not notice any difference among the participants, and the task was absolutely clear to everybody after the instructions and the practice.

Experiment 1: Effect of Speed—Elliptical Paths

The purpose of this experiment was to test the ability of observers to discriminate two similar relative motions and to assess how performance was affected by different types of common motion.

Method

Three types of common motion were used: translation, divergence, and rotation. The trajectory of the relative motion was either a circle (radius between 0.48° and 1.80° of visual angle) or an ellipse (main radius between 0.48° and 1.80° of visual angle) whose aspect ratio was always $1/2$. As mentioned in the General Method for Experiments 1–5 section, the relative motion could start at a randomized location along the contour of the circle or ellipse, and its speed was adjusted so that a single revolution was always completed. In other words, the speed of relative motion was higher for a larger circle than for a smaller one. The orientation of the main axis of the ellipse was randomly chosen in each trial.

The speed of the common motion was varied and had four levels: 1.4, 1.8, 2.4, and 3.6 deg/s. For translation, this is the speed of each element in the configuration. For divergence and rotation, speed varied with the location of the element in the region and stayed constant during the animation; the average was then computed based on the area of the region. Thus, in terms of velocities, the only difference between divergence and rotation was the direction of the motion.

The total number of unique conditions was 24 (4 levels of speed \times 3 types of common motion \times 2 types of relative motion), and each condition

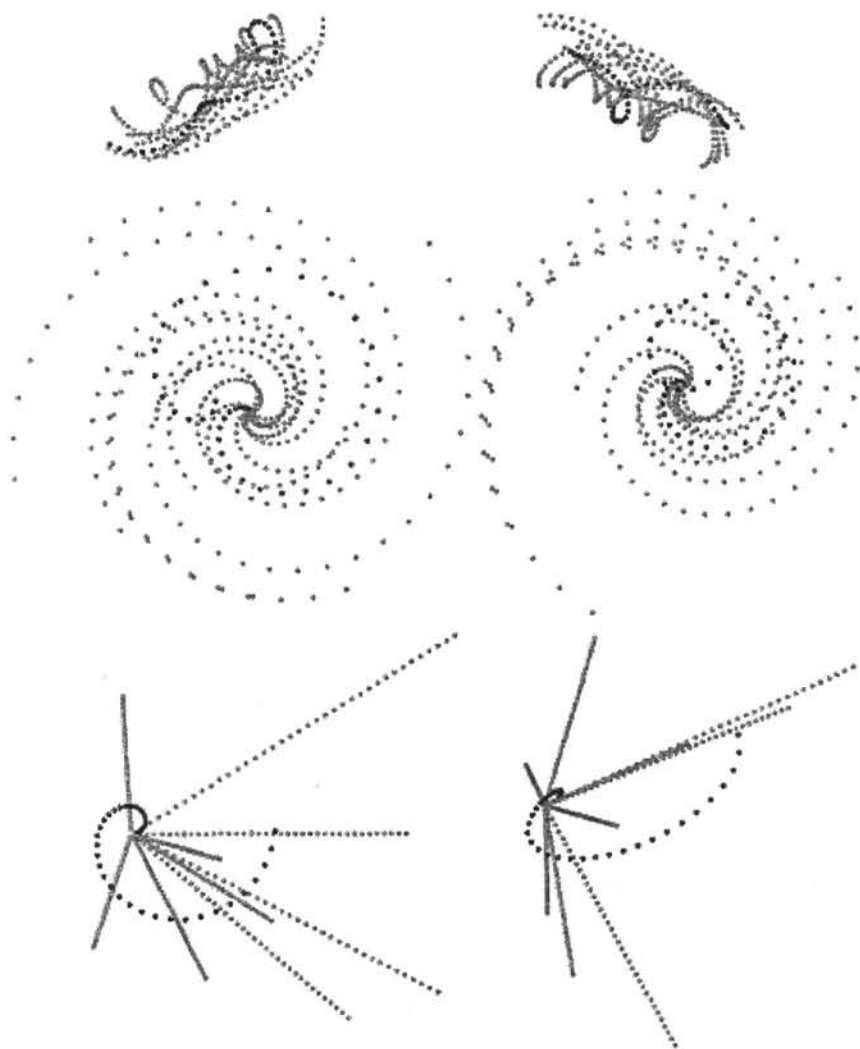


Figure 6. Trace of 10 dots during animation: examples randomly chosen from the set of displays for Experiment 4. The rows have three different types of common motion: translation + rotation, divergence + rotation, and divergence + translation. The relative motions are along a circle (on the right) or an ellipse (on the left).

was presented 15 times in a randomized order for a total of 360 trials. The trials were divided into five blocks of 72 trials, and an extra trial was added at the beginning of each block and excluded from the analysis. The 5 observers were encouraged to rest between blocks. The experiment lasted approximately 1 hr.

Results

The percentage correct was computed for each observer and was transformed in order to normalize the distribution. A nonlinear transformation called a folded root (Tukey, 1977) was used and is defined in Equation 1:

$$f(p) = \sqrt{2p} - \sqrt{2(1-p)} \quad (1)$$

The transformed data were subjected to a repeated measures analysis of variance (ANOVA) with common motion (three levels) and speed (four levels) as independent variables. Results are

shown in Figure 7A with within-subject confidence intervals (Lof-tus & Masson, 1995). These confidence intervals are consistent with the results from the ANOVA but are not the same as confidence intervals for population parameters because of the repeated measures design. Moreover, because these confidence intervals are relative to means for an interaction effect (Type of Motion \times Speed), they are based on the pooling of three error terms (omni-bus error term) under the assumption of sphericity. Note that, because of the transformation, the points are not arithmetic means of percentage correct data. Only the effect of type of motion was significant, $F(2, 8) = 31.949$, $p < .001$.

To compare sensitivity in the different conditions, we computed d' values for each observer based on correct or incorrect responses collapsed across speeds. Speed was not analyzed because of the limited number of data points but also on the basis of the analysis of the percentage correct data, where the effect of speed was not

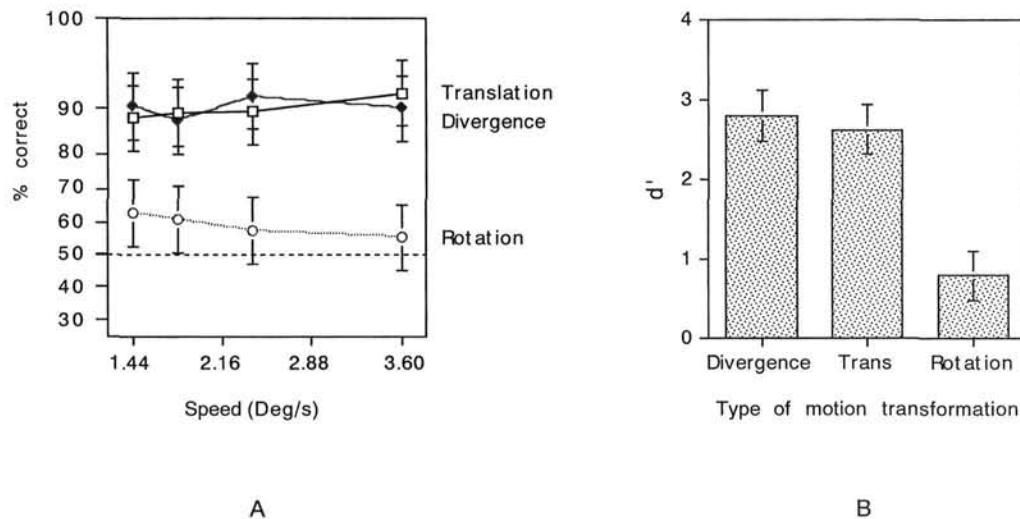


Figure 7. Data from Experiment 1. A: Percentage correct as a function of speed for different types of motion transformation. The percentage correct data have been transformed, and therefore the mean values are not arithmetic means of the raw data. The transformation is described in the text and is the same in all the graphs. The error bars are 95% within-subject confidence intervals for the means. B: d' for different types of motion. The error bars are 95% confidence intervals for the parameters. Trans = translation.

confirmed. Figure 7B shows the values for the d' (averaged across observers) for divergence, translation, and rotation. Standard errors of d' for each observer individually are computed according to the following formula (Macmillan & Creelman, 1991):

$$SE_{d'} = \sqrt{\frac{H(1-H)}{N[\phi(H)]^2}} + \sqrt{\frac{F(1-F)}{N[\phi(F)]^2}}, \quad (2)$$

where H and F are hit rate and false-alarm rate, N is the number of trials, and $\phi(x)$ is the height of the normal density function at $z(x)$. Confidence intervals were then based on the average standard error from the 5 observers, assuming independence. Both the collapsing of the data points across conditions and the assumption of independence may lead to some underestimation of the amount of variability.

Discussion

The analyses show that translation and divergence can provide a frame of reference because observers can detect different relative motion trajectories. Performance does not seem to depend on the speed of the common motion within the range of speeds used in this experiment. On the other hand, performance for rotation was not reliably above a chance level.

Experiment 2: Effect of Speed—Linear Paths

This experiment served as a control for the previous one. If the differences observed depend on the nature of the common motion, they should generalize to different types of relative motion trajectories. In this experiment, linear trajectories instead of elliptical trajectories were used.

Method

The stimuli were identical to those of Experiment 1, except for the relative motion trajectory that was either a square (side between 0.48°

and 1.80° of visual angle) or a rectangle (long side between 0.48° and 1.80° of visual angle, short side half of the long side). The relative motion could start at a randomized position along the trajectory, and its speed was adjusted so that a complete revolution (motion along the perimeter) was performed. The common motion had the following levels of average speed: 1.4, 1.8, 2.4, and 3.6 deg/s. This led to a total of 24 unique conditions presented 15 times (360 total trials) in a randomized order with four breaks. The procedure was identical to that of Experiment 1.

Results

Data were analyzed using the same steps as in the previous experiment. Percentage correct was computed for each observer and was transformed in order to normalize the distribution. The transformed data were subjected to a repeated measures ANOVA with common motion and speed as independent variables. Results are shown in Figure 8 with confidence intervals. Only the effect of type of motion was significant, $F(2, 8) = 15.848$, $p < .01$.

In a second ANOVA the results were compared with those of Experiment 2, and type of trajectory was a between-subjects variable. The independent variables were trajectory (elliptical vs. linear), type of motion (three levels), and speed (four levels). Performance for elliptical paths was confirmed to be significantly higher than for linear paths, $F(1, 8) = 10.707$, $p < .05$. The effect of type of motion was also significant, $F(2, 16) = 46.588$, $p < .001$, but there were no significant interactions.

To compare sensitivity in the different conditions, we computed d' values for each observer based on correct or incorrect responses collapsed across speeds following the same steps as in the analysis for the previous experiment. Speed was not analyzed because of the limited number of data points but also on the basis of the analysis of percentage correct data where the effect of speed was not confirmed. Figure 8 shows the values for the d' for divergence, translation, and rotation.

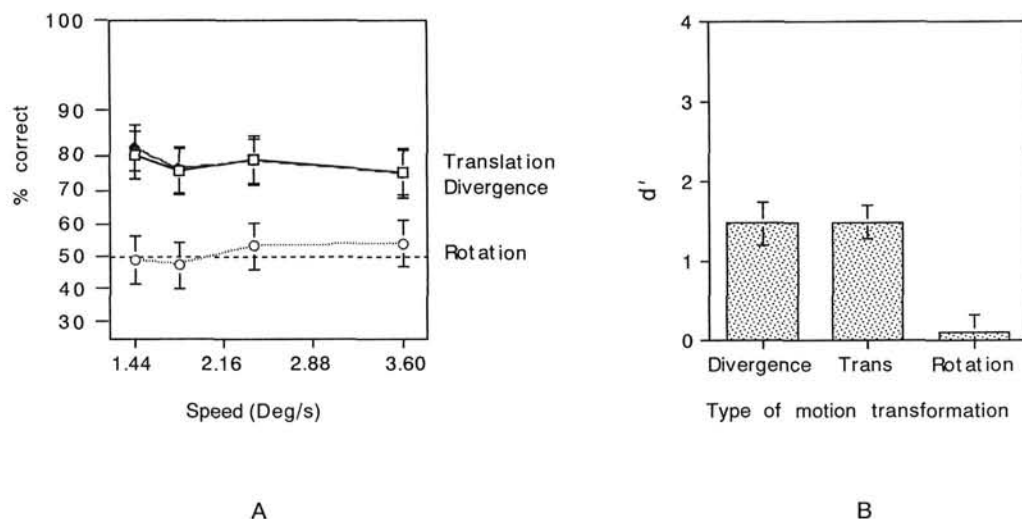


Figure 8. Data from Experiment 2. A: Percentage correct as a function of speed for different types of motion transformation. The error bars are 95% within-subject confidence intervals for the means. B: d' for different types of motion. The error bars are 95% confidence intervals for the parameters. Trans = translation.

Discussion

The analyses show that translation and divergence can provide a frame of reference because observers can detect different relative motion trajectories. On the other hand, performance for rotation was not above a chance level. This replicates the finding of Experiment 1. The only difference is that the level of performance was lower with linear paths. We have no simple explanation for that, but we find it interesting that the same pattern can be observed even with a task that appears to be more difficult.

Experiment 3: Effect of Density

This experiment tested the difference between types of motion for different densities of dots. If the differences observed in Experiments 1 and 2 depend on the nature of the common motion, they should generalize to other densities. However, fewer dots may fail to produce a strong enough frame of reference effect.

Method

The stimuli were similar to those of Experiment 1, except that the average speed was now constant at 1.8 deg/s. The dots were randomly located within the same circular region, but they could vary in number from 5 to 14 in steps of 3. This led to a total of 24 unique conditions (4 levels of density \times 3 types of common motion \times 2 types of relative motion) presented 15 times (360 total trials) in a randomized order with four breaks. The procedure was identical to that of Experiments 1 and 2.

Results

Data were analyzed using the same steps as in the previous two experiments. Percentage correct was computed for each observer and was transformed in order to normalize the distribution. The transformed data were subjected to a repeated measures ANOVA with common motion and speed as independent variables. Results are shown in Figure 9A with confidence intervals. Only the effect of type of motion was significant, $F(2, 8) = 22.585$, $p < .001$.

To compare sensitivity, we computed d' values for each observer based on correct or incorrect responses collapsed across densities. Density was not analyzed because of the limited number of data points but also on the basis of the analysis of percentage correct data where the effect of speed was not confirmed. Figure 9B shows the values for the d' for divergence, translation, and rotation.

Discussion

The analyses show that translation and divergence can provide a frame of reference. Performance for rotation was not systematically above a chance level. This replicates and extends the previous finding: The effect of type of motion seems to be the same for a range of dot densities. We did not use fewer than five dots, but it is conceivable that three dots would be too few to obtain this pattern of results, because it would be impossible to create a truly arbitrary distribution of dots with only three elements. This may explain the fact that our findings are only partly consistent with the results from Börjesson and Ahlström (1993). They used a configuration of five dots organized in two sets of three (with one dot shared) and found that rotation is a stronger grouping factor than divergence.

It is also possible that a much larger number of dots may lead to differences whenever the configuration starts to be perceived as a texture instead of a set of elements. We were interested, however, in this range of dots because this is what we think of as a configuration of elements: This number of dots is enough to create a whole but not enough to create a textured surface.

Experiment 4: Effect of Motion Combinations

In Experiments 1, 2, and 3, we compared performance for pure common motion transformations. In this experiment, transformations were combined to obtain more complex common motions. This distinguished between two possibilities. The first is that

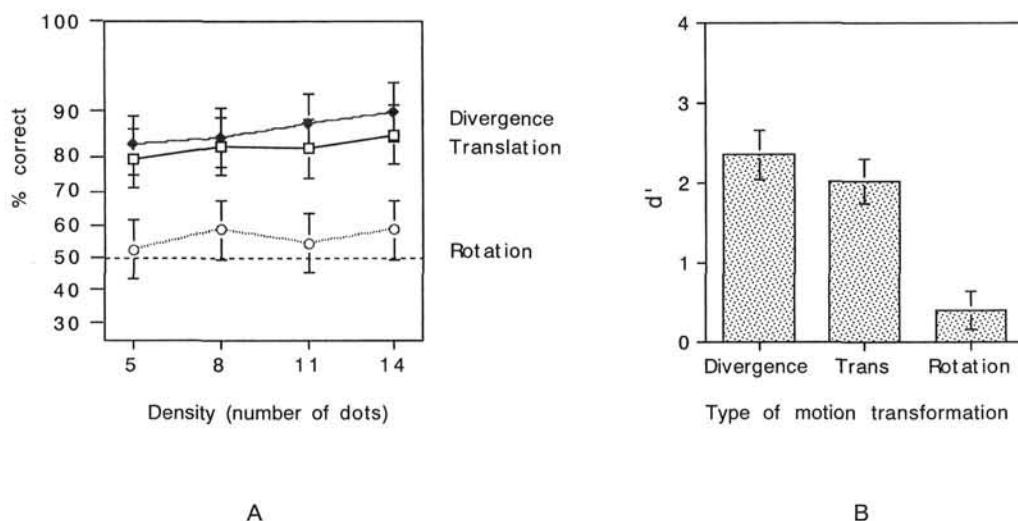


Figure 9. Data from Experiment 3. A: Percentage correct as a function of speed for different types of motion transformation. The error bars are 95% within-subject confidence intervals for the means. B: d' for different types of motion. The error bars are 95% confidence intervals for the parameters. Trans = translation.

translation and divergence may always create groupings and, therefore, hierarchical frame of reference effects. If so, people will be able to see the hierarchical motion separation even when a rotation is added to the common translation or divergence. The second possibility, on the other hand, is that the poor performance for rotation is due to something intrinsic in a rotary transformation, and whenever it is present, people will not extract a relative motion trajectory.

Method

The stimuli were similar to those of Experiment 1, except for the common motion that was always a combination of two motions: translation + divergence, translation + rotation, or rotation + divergence. The relative motion trajectory was either a circle or an ellipse. The common motion had the following levels of average speed: 1.4, 1.8, 2.4, and 3.6 deg/s. This led to a total of 24 unique conditions (4 levels of speed \times 3 common motion combinations \times 2 types of relative motion) presented 15 times (360 total trials) in a randomized order with four breaks. The procedure was identical to that of Experiment 1.

Results

Data were analyzed using the same steps as in the previous experiments. Percentage correct was computed for each observer and was transformed in order to normalize the distribution. The transformed data were subjected to a repeated measures ANOVA with common motion and speed as independent variables. Results are shown in Figure 10 with confidence intervals. Only the effect of motion combinations was significant, $F(2, 8) = 20.622$, $p < .001$.

To compare sensitivity, we computed d' values for each observer based on correct or incorrect responses collapsed across densities. Speed was not analyzed because of the limited number of data points but also on the basis of the analysis of percentage correct data where the effect of speed was not confirmed. Fig-

ure 10 shows the values for the d' for divergence, translation, and rotation.

Discussion

The analyses show that translation + divergence can provide a frame of reference because observers can detect different relative motion trajectories. On the other hand, performance for translation + rotation and divergence + rotation was not systematically above a chance level. This is consistent with the idea that whenever a rotation is present, the task becomes very difficult. It is not consistent, however, with the idea that translation and divergence transformations are sufficient to create a frame of reference. This result may lead to a revision of the basic hypothesis. We stressed that translation and divergence may be important because they provide information about where the object is going (translation in space). It may be that what should be stressed more is the fact that people find it difficult to see rotating configurations as grouped with hierarchically organized motions. That is, rotation may be the culprit for the poor performance instead of translation and divergence being responsible for the good performance.

We have cast our conclusions with respect to the types of common motion. It is important to observe that alternative terminologies are possible—for instance, one that stresses which aspects of the paths of motion presented correlate with poor performance. Experiment 4 led us to conclude that there is something about rotation that makes the task difficult. As pointed out before, rotation defines a center, it is periodic, and the path of each individual dot is curved. The presence of a center around which speed is a linear function of eccentricity is in common with divergence and therefore cannot be the problem (this makes even straight paths of relative motion curved with respect to the monitor in both cases). The periodicity is unlikely to be the problem because in all our experiments we never go farther than one full period. The only thing left is that the paths of the common motion

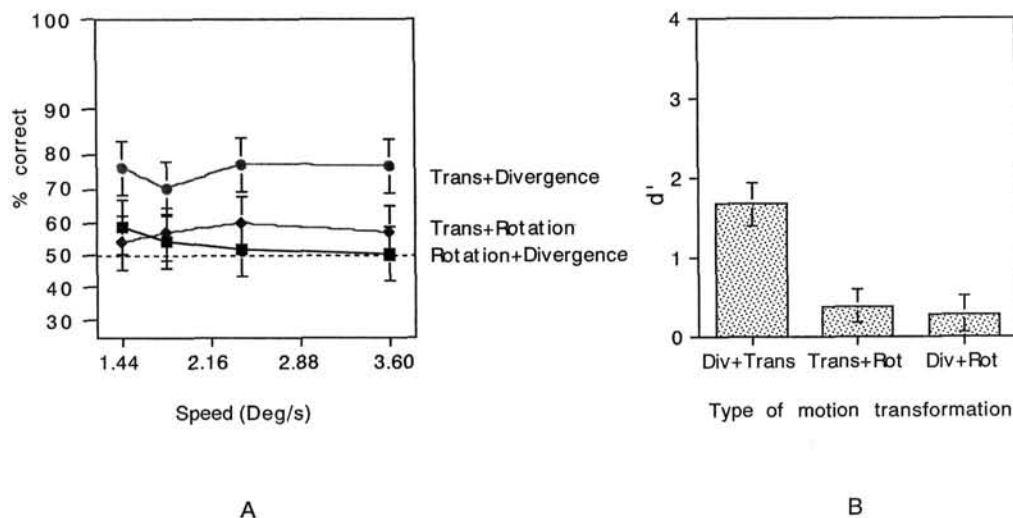


Figure 10. Data from Experiment 4. A: Percentage correct as a function of speed for different types of motion transformation. The error bars are 95% within-subject confidence intervals for the means. B: d' for different types of motion. The error bars are 95% confidence intervals for the parameters. Div = divergence; Trans = translation; Rot = rotation.

themselves are curved, unlike in the other two conditions. In other words, in a pure rotation the observer needs to update the orientation of the frame of reference continuously.

Experiment 5: Effect of SI Constraints

This experiment deals with nonrandom configurations in the following sense. The relative motion of the dot is not simply superimposed (added) onto the background motion: instead, the combination of the two motions is constrained in such a way that an SI is preserved. The relative circular motion is now always centered on one of the visible elements of the background. As a consequence, in the translation and rotation conditions, the radius of the circular motion will be the fixed distance between two of the dots in the display. In the divergence condition this radius will instead increase linearly over time. For the trials in which the relative motion is elliptical, the visible elements were constrained to be in the center of the ellipse (midpoint between the two loci). We predicted that the dot at the center of the relative motion would be seen as the pivot point for the relative trajectory of the dot and, therefore, would allow a discrimination between the circular and the elliptical motions.

It is important to note that the presence of SIs does not alter the motion information, which is identical to the motion of Experiment 1. Thus, any difference in the findings must be due to the SI constraint.

Method

The stimuli were identical to those of Experiment 1, except that the relative motion was constrained to be centered on one of the dots of the configuration undergoing the common motion. Speed was varied as in Experiment 2; therefore, there were 24 unique conditions (4 levels of speed \times 3 types of common motion \times 2 types of relative motion) presented 15 times (360 total trials) in a randomized order with four breaks. The procedure was identical to that of Experiment 1.

Results

Data were analyzed using the same steps as in the previous experiments. Percentage correct was computed for each observer and was transformed in order to normalize the distribution. The transformed data were subjected to a repeated measures ANOVA with common motion and speed as independent variables. Results are shown in Figure 11 with confidence intervals. None of the effects were significant; in particular, there was no difference between types of motion, $F(2, 8) = 3.464$, *ns*.

In a second ANOVA, which introduced SI as a between-subjects variable, performance in Experiments 1 and 5 was confirmed to be significantly different. The independent variables were SI (present vs. absent), type of motion (three levels), and speed (four levels). The effect of type of motion was significant, $F(2, 16) = 17.653$, $p < .001$, and the interaction of type of motion with SI was also significant, $F(2, 16) = 5.741$, $p < .05$.

To compare sensitivity, we computed d' values for each observer based on correct or incorrect responses collapsed across speeds. Speed was not analyzed because of the limited number of data points but also on the basis of the analysis of correct responses where the effect of speed was not confirmed. Figure 9 shows the values for the d' for divergence, translation, and rotation.

Discussion

Figure 11 should be compared with Figure 7. When an SI constraint is introduced, people can exploit this information to perform the discrimination task and thereby perform at a level of approximately 75% correct even for rotation. On the other hand, it seems that performance for divergence, which had been above that for translation, is now lower. Both of these findings can be interpreted as an effect of the SI that was introduced. The fixed radius is present for translation and rotation, but this distance changes with divergence because of the rescaling of the whole configura-

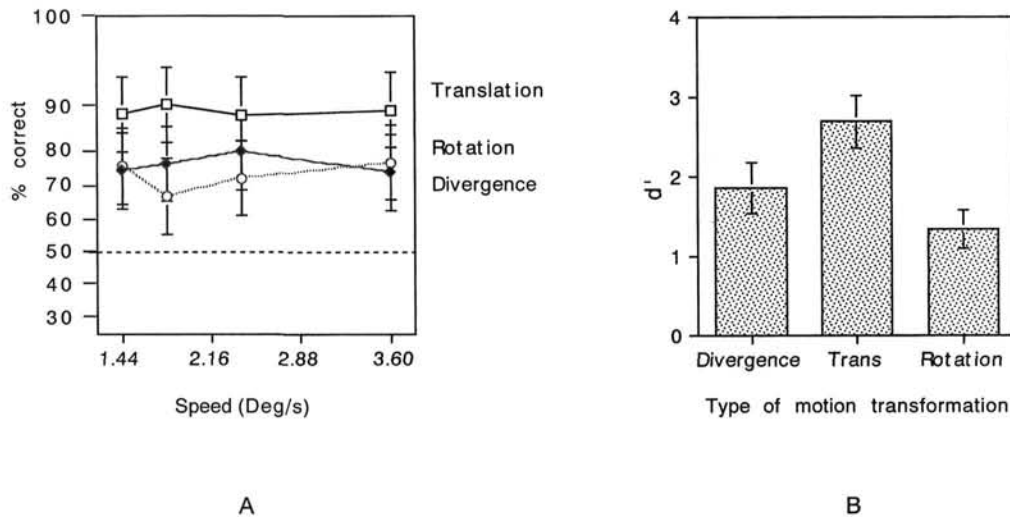


Figure 11. Data from Experiment 5. A: Percentage correct as a function of speed for different types of motion transformation. The error bars are 95% within-subject confidence intervals for the means. B: d' for different types of motion. The error bars are 95% confidence intervals for the parameters. Trans = translation.

tion. This finding is particularly important because the dynamic information was identical in Experiments 1 and 5. The implication is that hierarchical motion organization also depends on SI information and that it is this kind of information that can allow people to perform tasks that are otherwise very difficult such as when the common motion is a rotation.

Why should SI information be important? Most of the objects in our terrestrial environment have constraints on their motions that come from their physical structure. A tree is a hierarchy of furcations; the body of a human, and that of most animals, is a hierarchy of limbs; and even a falling object will spin around a center of gravity that in many cases corresponds to its centroid. Motion needs to be organized even in cases in which—because of noise in the input due, for example, to multiple occlusion—only parts of an object are visible. Visible parts are in general not arbitrary; instead the likelihood of their locations depends on the structure of the underlying object.³

Another way to think about SI information is to compare it to the assumption of pairwise rigid planar motions (Hoffman & Flinchbaugh, 1982) or fixed-axis motions (Hoffman & Bennett, 1986). The first assumption has been introduced in the context of biological motion to explain the level of performance in recognizing the motion of terrestrial bipeds and quadrupeds. More specifically, the assumption that two points are rigidly connected (a fixed axis) and rotating in a plane makes the structure recoverable from three orthographic projections. Three points forming two hinged pairs are recoverable from only two projections (Hoffman & Flinchbaugh, 1982). A fixed-axis assumption can also be applied to elements rotating at varying angular velocities; this structure can be recovered from three projections of four elements sharing the same axis (Hoffman & Bennett, 1986). In a sense, the idea of SI in Experiment 5 is complementary to these assumptions. Instead of assuming fixed axes and coplanarity to constrain motion interpretations, we suggest that when there is evidence of a fixed axis in the animation, it immediately informs the process of hierarchical

motion interpretation. In other words, in our experiments fixed axes are not assumed, they are deduced from the evidence coming from SI.

Experiment 6: Target-Dot Visibility

This last experiment deals with the issue of visibility of the dot with relative motion. It is quite possible that the effect of a rotation is to make it more difficult to find the dot that moves with a relative motion, simply because more time is necessary to locate this target dot. On the other hand, once such a target dot is found, then performance could be similar in the relative motion discrimination task. To test this hypothesis, we conducted a new experiment that replicates the conditions of Experiment 1, except for one important modification. The dot with relative motion is now red, whereas all the other dots are black. Therefore, the visibility of the target dot is high and there is absolutely no ambiguity about which dot is carrying the relative motion.

Method

The stimuli were similar to those of Experiment 1 except that the relative motion was carried by a red dot whereas the other nine dots were black on a white background. For practical reasons, the experiment had to be ported to the Macintosh platform, using a PowerComputing PowerTowerPro 225 driving a Sony Trinitron 17sffII monitor (at 75 Hz). The specifications in the General Method for Experiments 1–5 section still apply to Experiment 6, except for the differences mentioned in this section.

³ Interestingly, animals who camouflage (zebras, leopards, and so on) have uniform patterns of lines or dots, so that the pattern does not have any relation with the structure of their body. On the other hand, some reef fish that need to escape predators have dots resembling eyes at the rear end of their body so that this cue is misleading as to the direction of motion of the fish.

Speed was varied as in Experiment 1; therefore, there were 24 unique conditions (4 levels of speed \times 3 types of common motion \times types of relative motion) presented 15 times (360 total trials) in a randomized order with four breaks. Care was taken to keep all parameters identical to those of Experiment 1. The values of speed were different (1.8, 2.7, 3.6, and 4.5 deg/s) because of the different hardware, but they largely overlap with the values of Experiment 1 (1.44, 2.16, 2.88, and 3.6 deg/s).

Results

Data were analyzed using the same steps as in the previous experiments. Percentage correct was computed for each observer and was transformed in order to normalize the distribution. The transformed data were subjected to a repeated measures ANOVA with common motion and speed as independent variables. Results are shown in Figure 12 with confidence intervals. Only the effect of type of motion was significant, $F(2, 8) = 6.745$, $p < .02$.

To compare, we computed sensitivity d' values for each observer based on correct or incorrect responses collapsed across speeds. Speed was not analyzed because of the limited number of data points but also on the basis of the analysis of correct responses where the effect of speed was not confirmed. Figure 10 shows the values for the d' for divergence, translation, and rotation.

Discussion

Figure 12 should be compared with Figure 7. Performance in Experiment 6 was lower, but in terms of the relative effect of the motion types, the differences between the two plots are small, and overall we can consider Experiment 6 as a replication of the finding of Experiment 1. Namely, the task is more difficult for rotation than for translation and divergence, whereas performance levels for translation and divergence are similar. The lower overall level of performance may be linked perhaps to a less effective procedure of familiarization and practice or to the fact that the room was not quite as dark as the room used in the previous

experiments. The higher illumination makes the monitor more visible, and this physical frame probably competes with the moving set of dots as a frame of reference.

The implication of Experiment 6 is that hierarchical motion organization is more difficult for rotation not as a consequence of a diminished visibility of the dot carrying the relative motion.

General Discussion

In our experiments, we found that translation and divergence can easily provide a moving frame of reference, whereas rotation cannot. The task in our studies was to discriminate two similar relative motion trajectories within a configuration of moving dots. Performance required a perceptual organization of the displays into the two motion components, a common motion of the configuration as a whole and a relative motion of a single dot. The number of dots in the configuration (density) and speed were varied, and the basic finding held within the range of values manipulated for these variables. Thus, we can discard the second and third hypotheses we set forth in the introduction—namely, that all motion transformations are equivalent or, alternatively, that translation is superior to all other transformations.

The fact that performance for the divergence conditions was as good as for the translation ones suggests that we cannot explain the finding as a consequence of a more basic nature (mathematically or physiologically) of translation. Moreover, if the priority of transitory motion derived from its mathematical simplicity or its physiological status, we would expect this advantage to show up in other types of human performance. Price and Gilden (2000) found that memory tasks are affected by these different motion transformations in the same manner as was found in our perceptual organization experiments. They found that translation, as well as divergence, produces high recall for both direction of motion and

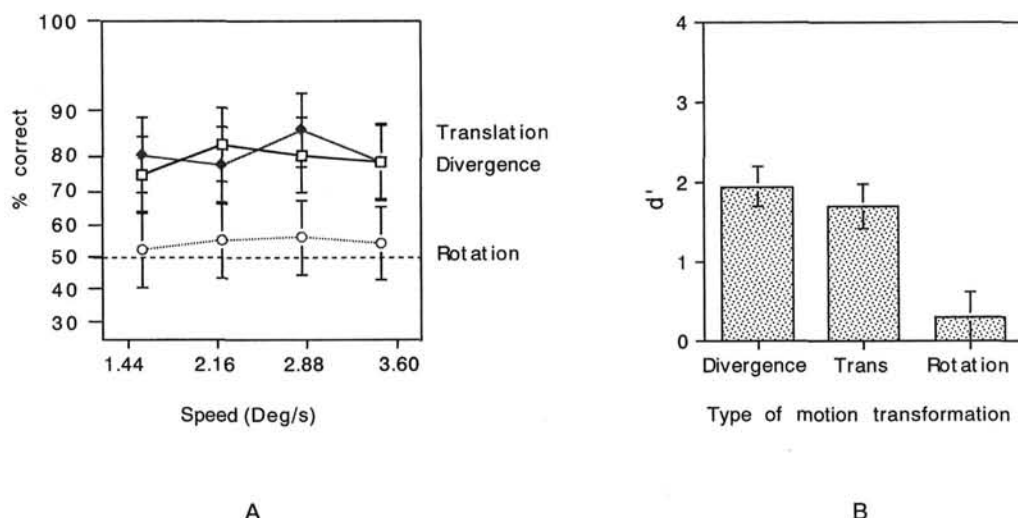


Figure 12. Data from Experiment 6. A: Percentage correct as a function of speed for different types of motion transformation. The error bars are 95% within-subject confidence intervals for the means. B: d' for different types of motion. The error bars are 95% confidence intervals for the parameters.

speed, whereas memory for rotation was at a chance level.⁴ The fact that performance for divergence was also very high suggests that the critical variable is not the fact that translation can be processed in parallel due to its local nature; divergence in fact requires the location of a center of expansion just like rotary motion requires the location of a center of rotation. It is possible that the critical variable is the absence in rotation of an overall displacement of the region. Our findings also imply that the possibility of tracking a pure translation is not necessary for performing this task. Divergence and rotation conditions were matched not only in terms of average speed but in terms of the set of speeds present in the display (in both cases, speed varied as a function of location or eccentricity, and there was no acceleration in either condition). Therefore, it seems that something else makes rotation an ineffective frame of reference.

The fact that pure isometric transformations do not contain any deformation or relative motion does not mean that there is no information from motion parallax. An arbitrary translatory motion of a set of elements can produce something close to pure translation and divergence only when all the elements belong to a single frontoparallel plane and, therefore, are close in space and rigidly connected. If differences in depth are present, they would produce motion parallax. The same is not true for rotation. A rotation of the field due to rotation of the head (around the line of sight) or a rotation of an object (around the line of sight) will be very close to a pure rotation because elements at different distances from the observer will have an instantaneous velocity that depends on the angular velocity of the rotation and the distance from the axis of rotation, but not on the distance from the observer (Simpson, 1993). In terms of nonaccidental properties, a common translation or divergence is unlikely to be produced by something other than a single compact object (i.e., the elements must be roughly coplanar). A common rotation is also unlikely to be produced by anything that is not rigidly connected, ignoring for the moment the problem of head or eye rotation, but it will be produced even if there are depth differences between elements of the object (i.e., the elements are not coplanar). This point can be clarified by considering the following. The probability that noncoplanar moving elements would produce a pure translation or a pure divergence is virtually zero. The probability that noncoplanar moving elements would produce a pure rotation is one for any set of rigidly connected elements rotating around the line of sight. This simple analysis is complicated only by the fact that pure rotations can also be produced by self-motion, although this does not change the main point that pure rotations do not signal coplanarity.

One possibility, therefore, is that pure rotation does not specify a coplanar surface the way that translation and divergence do. This may be a link between motion parallax and hierarchical motion organization. Observers may be able to extract relative motion within a coplanar configuration better than within a configuration with ambiguous depth relations. On the other hand, an alternative interpretation of this finding is that observers have a specific difficulty extracting and using rotary motion as a moving frame of reference. Rotation is special because it is periodic (even though only one revolution was presented in all our displays) and because it introduces a continuous change of orientation with respect to the horizontal-vertical axes of the picture plane. This alternative view is supported by the results of Experiment 4, in which combinations of rotation with translation and rotation with divergence led to

poor performance. Note that this change of orientation-direction specific to rotation always produces a curved path for the target dot, even when the relative motion of the target is a linear path, as in Experiment 2. This is true also for divergence, except that in that case the rest of the background dots were still moving in straight paths.

The prime role of translation is consistent with the results of Börjesson and Ahlström (1993), but the rest of our findings are not—in particular, they found that performance for rotation was good. Apart from the problem with their methodology described in the introduction, one reason may have to do with the fact that Börjesson and Ahlström used few dots (two groups of three) and they were not located randomly in a region but formed regular constellations. These constellations may create special arrangements and SIs. This factor was manipulated in Experiment 5, and the results show that the introduction of an SI constraint in a display changes the difficulty of the task. Performance for rotation in Experiment 5 was higher, because observers could base their judgments on a different source of information, and performance for divergence was lower (with respect to translation) because the SIs are less useful as a consequence of the scaling of the configuration. The fixed radius of the relative rotation was in this case not constant but increasing linearly.

Conclusions

When one is perceiving moving objects, it is necessary both to recognize where they are going and to organize their internal behavior so as to recognize what they are. This goal is realized through the perception of moving frames of reference. We compared different kinds of common motions and found large differences in the degree to which they could support moving reference frames. In Experiments 1–3, it was found that both translation and divergence can serve as moving frames of reference, whereas rotation cannot. Experiment 4 showed that it is the presence of rotation that makes the task of hierarchical motion organization difficult, because when rotation was combined with translation or divergence, performance fell to a chance level. The finding that rotation made the perceptual motion organization task very difficult holds for a range of speeds and densities.

A second factor was also found to constrain motion organization. SIs are constant spatial relationships revealed through motion, such as alignment and pivoting. Experiment 5 shows that the presence of SIs strongly affects motion organization. In Experiment 5, observers performed at levels well above chance, even for rotation. This is an effect of structure on motion, and it is interesting to note that attempts to integrate motion and shape information have begun in recent modeling of the visual system (Weiss & Adelson, 1994).

⁴ Because of the similarity of the results, one can ask the question of whether our results are caused by a memory effect. This hinges on how strictly one defines *memory*; in all our experiments the response was collected within a few seconds of the onset of the stimulus, and the decision was a simple two-alternative forced choice. It seems more likely that something about encoding the information of a rotating frame of reference is the cause of the effect, and perhaps this affects every kind of memory task.

An important implication of these findings is that a purely geometrical approach is insufficient to describe people's performance in hierarchical motion organization. Translation can be said to be simpler than divergence and rotation; moreover, because translation is locally specified, many researchers believe that translation is processed earlier by the visual system (e.g., Kosslyn & Andersen, 1992). Be that as it may, translation and divergence are roughly equivalent in their ability to support moving reference frames, whereas rotations cannot support this function without the presence of SIs. An ecological point of view makes better sense of these findings because both translation and divergence signify the displacement of a compact object whereas rotation does not.

The other conclusion of our studies is that SI information is important in determining what motion we perceive. This supports the view that motion and spatial structure are not really visually separate.

References

- Adelson, E. H., & Bergen, J. R. (1985). Spatiotemporal energy models for the perception of motion. *Journal of the Optical Society of America*, 2, 284–299.
- Börjesson, E., & Ahlström, U. (1993). Motion structure in five-dot patterns as a determinant of perceptual grouping. *Perception & Psychophysics*, 53, 2–12.
- Börjesson, E., & von Hofsten, C. (1972). Visual perception of motion in depth: Application of vector model to three-dot motion patterns. *Perception & Psychophysics*, 13, 169–179.
- Börjesson, E., & von Hofsten, C. (1975). A vector model for perceived object rotation and translation in space. *Psychological Research*, 13, 169–179.
- Cutting, J. E., & Proffitt, D. R. (1982). The minimum principle and the perception of absolute, common, and relative motions. *Cognitive Psychology*, 14, 211–246.
- Dunker, K. (1937). Induced motion. In W. D. Ellis (Ed. and Trans.), *Source book of Gestalt psychology* (pp. 161–172). London: Routledge & Kegan Paul. (Original work published 1929 in German)
- Eizenman, D. R., & Bertenthal, B. I. (1998). Infants' perception of object unity in translating and rotating displays. *Developmental Psychology*, 34, 426–434.
- Gogel, W. C. (1978). The adjacency principle in visual perception. *Scientific American*, 238(5), 126–139.
- Hochberg, J. (1986). Representation of motion and space in video and cinematic displays. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance* (Vol. 1, pp. 1–64). New York: Wiley.
- Hoffman, D. D., & Bennett, B. M. (1986). The computation of structure from fixed-axis motion: Rigid structures. *Biological Cybernetics*, 54, 71–83.
- Hoffman, D. D., & Flinchbaugh, B. E. (1982). The interpretation of biological motion. *Biological Cybernetics*, 42, 195–204.
- Johansson, G. (1950). *Configurations in event perception*. Uppsala, Sweden: Almqvist & Wiksell.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14, 201–211.
- Johansson, G. (1977). Studies on visual perception of locomotion. *Perception*, 6, 365–376.
- Kappers, A. M. L., te Pas, S. F., Koenderink, J. J., & van Doorn, A. J. (1996). Simulating the detection of first-order optical flow components. *Vision Research*, 36, 3539–3547.
- Kellman, J. K., & Spelke, E. S. (1983). Perception of partly occluded objects in infancy. *Cognitive Psychology*, 15, 483–524.
- Koenderink, J. J. (1986). Optic flow. *Vision Research*, 26, 161–180.
- Köhler, W. (1947). *Gestalt psychology: An introduction to new concepts in modern psychology*. New York: Liveright.
- Kosslyn, S. M., & Andersen, R. A. (1992). *Frontiers of cognitive neuroscience*. Cambridge, MA: MIT Press.
- Lappin, J. S., Norman, J. F., & Mowafy, L. (1991). The detectability of geometric structure in rapidly changing optical patterns. *Perception*, 20, 513–528.
- Loftus, G. R., & Masson, M. E. J. (1995). Using confidence intervals in within-subjects designs. *Psychonomic Bulletin & Review*, 1, 476–490.
- Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. Cambridge, England: Cambridge University Press.
- Morrone, M. C., Burr, D. C., & Vaina, L. M. (1995). Two stages of visual processing for radial and circular motion. *Nature*, 376, 507–509.
- Orban, G. A. (1992). The analysis of motion signals and the question of the nature of processing in the primate visual system. In G. A. Orban & H.-H. Nagel (Eds.), *Artificial and biological visual systems* (pp. 24–57). Berlin: Springer.
- Pittenger, J. B., & Shaw, R. E. (1975). Aging faces as visco-elastic events: Implications for a theory of nonrigid shape perception. *Journal of Experimental Psychology: Human Perception and Performance*, 1, 374–382.
- Price, C. M., & Gilden, D. L. (2000). Representations of motion and direction. *Journal of Experimental Psychology: Human Perception and Performance*, 26, 18–30.
- Proffitt, D. R., Cutting, J. E., & Stier, D. M. (1979). Perception of wheel-generated motions. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 289–302.
- Reichardt, W. (1961). Auto-correlation: A principle for the evaluation of sensory information by the central nervous system. In W. A. Rosenblith (Ed.), *Sensory communication* (pp. 303–317). New York: Wiley.
- Restle, F. (1979). Coding theory and the perception of motion configurations. *Psychological Review*, 86, 1–24.
- Rubin, E. (1927). Visuell wahrgenommene wirkliche Bewegungen [Visually perceived real motions]. *Zeitschrift für Psychologie*, 103, 384–392.
- Shepard, R. N., & Zare, S. (1983). Path-guided apparent motion. *Science*, 220(4597), 632–634.
- Shum, K. H., & Wolford, G. L. (1983). A quantitative study of perceptual vector analysis. *Perception & Psychophysics*, 34, 17–24.
- Simpson, W. A. (1993). Optic flow and depth perception. *Spatial Vision*, 7, 35–75.
- Tukey, J. W. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.
- van Santen, J. P. H., & Sperling, G. (1985). Elaborated Reichardt detectors. *Journal of the Optical Society of America A*, 2, 300–321.
- Weiss, Y., & Adelson, E. H. (1994). *Perceptually organized EM: A framework for motion segmentation that combines information about form and motion* (Tech. Rep. No. 315). Cambridge, MA: MIT Media Laboratory.
- Wertheimer, M. (1937). Laws of organization in perceptual forms. In W. D. Ellis (Ed. and Trans.), *A source book of Gestalt psychology* (pp. 71–88). London: Routledge & Kegan Paul. (Original work published 1923 in German)

Received July 17, 1997

Revision received June 18, 1999

Accepted August 5, 1999 ■