

The perception of biological motion across apertures

MAGGIE SHIFFRAR, LAURA LICHTY, and SHEBA HEPTULLA CHATTERJEE
Rutgers University, Newark, New Jersey

To understand the visual analysis of biological motion, subjects viewed dynamic, stick figure renditions of a walker, car, or scissors through apertures. As a result of the aperture problem, the motion of each visible edge was ambiguous. Subjects readily identified the human figure but were unable to identify the car or scissors through invisible apertures. Recognition was orientation specific and robust across a range of stimulus durations, and it benefited from limb orientation cues. The results support the theory that the visual system performs spatially global analyses to interpret biological motion displays.

How does the visual system analyze patterns of movement generated by walking animals? Johansson (1973) initiated the systematic study of “biological motion” perception by demonstrating that observers could readily recognize simplified depictions of human locomotion. In a darkened environment, Johansson and his colleagues filmed the movements of individuals with point light sources attached to their major joints. Observers of the films were rapidly able to identify the movements generated by the point-light-defined actors even though the displays were nearly devoid of form information. Subsequent research indicates that our ability to recognize point light walkers is rapid (Johansson, 1976), develops as early as 3 months of age (Bertenthal, Proffitt, Kramer, & Spetner, 1987), and extends to the recognition of friends (Cutting & Kozlowski, 1977) and other animals (Mather & West, 1993).

One of the central issues in the study of biological motion perception concerns whether the primary level of analysis is local or global. While *local* and *global* are difficult to define as absolute terms, most researchers in this field define local analyses as the computations conducted within small spatiotemporal neighborhoods on individual points (joints) or point pairs (limbs). Global analyses are conducted over larger areas and generally involve at the very least two limbs up to an entire point light walker. Early models of point light walker perception emphasized the local detection of pairs of rigidly connected points. These models took advantage of the fact that pairs of points corresponding to the joints at either end of a limb are rigidly connected. The identifica-

tion of rigidly connected point pairs was followed by a combination of the pairs into larger groups until a correctly connected human form was constructed (Hoffman & Flinchbaugh, 1982; Webb & Aggarwal, 1982). However, these models are unable to explain the orientation specificity of biological motion perception. That is, although observers can readily identify an upright point light walker, recognition rates fall substantially when point light walkers are presented upside-down (see, e.g., Bertenthal, 1993; Sumi, 1984).

Behavioral evidence supporting the global analysis of biological motion usually involves masked point light walker displays. Masks are thought to reduce the accuracy of local analyses by surrounding each walker point with other moving points to increase the number of possible false correspondences. Studies based on this approach have demonstrated that subjects are able to detect the presence as well as the direction of motion of upright point light walkers in random dot masks (Cutting, Moore, & Morrison, 1988) and scrambled limb masks (Bertenthal & Pinto, 1994). The global analysis of biological motion is also suggested from neurophysiological studies of the macaque temporal lobe. Neurons in the anterior region of the superior temporal polysensory area (STPa) respond to precise combinations of biological motions and forms but not to inanimate control objects (Perrett, Harries, Mistlin, & Chitty, 1990). For example, an STPa neuron might respond selectively to a forearm extending outward from the elbow but not to that forearm when it contracts or to a rotating bar that replicates the forearm's motion. Moreover, while these STPa neurons respond vigorously to entire point light walker displays, they remain unresponsive to partial displays (Oram & Perrett, 1994).

If biological motion plays a fundamental role in our interactions with the environment, perceptual sensitivity to biological motion should not be limited to the analysis of point light walker displays. In the present set of experiments, we examined whether the global analysis of biological motion would extend to the integration of ve-

This work was funded by NIH Grant 099310 to the first author. Some of these results were presented at the 1994 Cognitive Neuroscience Conference in San Francisco. We thank Bennett Bertenthal, Judy Kegl, and an anonymous reviewer for numerous helpful suggestions on a previous draft of this manuscript. We also thank Jean Dominique LaJoux for generously providing valuable information on the research of E. J. Marey. Correspondence should be addressed to M. Shiffrar, Center for Neuroscience, 197 University Avenue, Rutgers University, Newark, NJ 07102 (e-mail: mag@cortex.rutgers.edu).

locity estimates across disconnected contours. That is, while point light walker displays provide information about the changing positions of a walker's joints, we examined how the visual system interprets the motion of a walker's limbs. We adapted a technique first devised by Étienne-Jules Marey (1895/1972) in which he measured human movement by placing a thin, relatively bright band along the length of each limb of a walker. Marey then photographed the walker so that only the bands were visible. We modified this technique in order to study the analysis of biological motion across disconnected line segments. We also conducted a preliminary test of the hypothesis that biological motion is analyzed more globally than nonbiological motions.

Our approach involved the creation of stimuli having different local and global interpretations. The aperture problem can be used to examine spatially separated local and global motion analyses (Shiffrar & Pavel, 1991; Shimojo, Silverman, & Nakayama, 1989). When a moving line is viewed through a relatively small window or aperture, the motion of the line is ambiguous because the component of translation parallel to the line's orientation cannot be measured. As a result, the line's motion is consistent with an infinitely large family of different motion interpretations, as shown in Figure 1.

The visual system can overcome this motion ambiguity through local or global levels of analysis. At the local level, the visual system can uniquely interpret the ambiguous stimulus by relying on the motion of discontinuities available within the aperture. In a classic series of experiments, Wallach (1935, 1976) found that even though the motion of a line translating behind an aperture is ambiguous,

observers often interpret the line to translate in the direction of its terminators. The best known example of this analysis is the barber pole illusion shown in Figure 1. Thus, a local solution of the aperture problem is to use the terminator motion to disambiguate the line motion. A global solution to the aperture problem includes linking motion signals across spatially disconnected image regions or apertures. That is, while the motion measurement of a single translating line is consistent with infinitely many interpretations, multiple measurements of differently oriented lines can be combined (by the intersection of constraints or a vector average) in order to interpret the motion uniquely (Adelson & Movshon, 1982; Mingolla, Todd, & Norman, 1992; Wilson, Ferrera, & Yo, 1992).

In previous studies, researchers have investigated how observers interpret moving stimuli having different local and global interpretations. In one series of studies, a simple translating or rotating square figure was viewed through multiple apertures positioned over the square's straight edges. A local analysis of the stimulus would lead to the interpretation of each visible edge moving independently in the direction defined by its terminators. A global analysis of this stimulus would involve the combination of motion signals across the edges and lead to the interpretation of a rigidly moving square. Behavioral studies demonstrate that under numerous conditions, observers interpret this stimulus as four independently moving edges rather than a coherently moving square. This finding is impressive, because observers do not combine motion signals across disconnected edges even when they have prior knowledge of the shape of the underlying figure. Thus, when local and global analyses lead to dif-

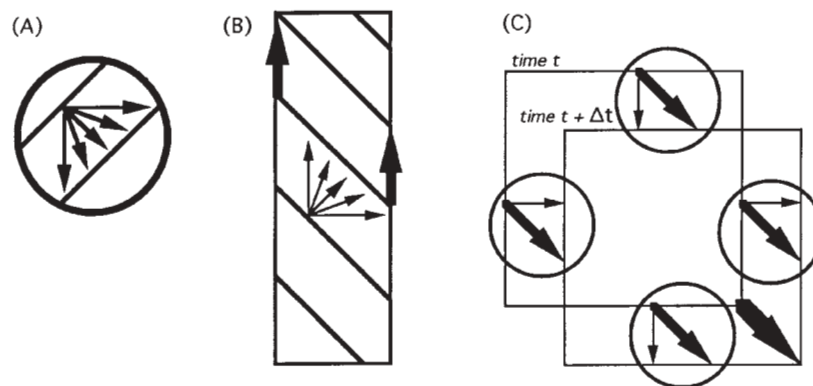


Figure 1. (A) The aperture problem. When a moving unmarked edge is viewed through a relatively small aperture, the motion parallel to the orientation of the line cannot be measured, because all the points along the line are identical. As a result, an entire family of motions with the same component of motion perpendicular to the orientation of the line but a differing parallel component of motion will appear to be identical. (B) The barber pole illusion. Although the motion of the lines is ambiguous, observers perceive vertical, upward motion when they rely on the terminators. (C) A square translating downward and to the right, seen at two different times. The square is viewed through four apertures, which are represented here as circles. As a result of the aperture problem, the motion measured within each receptive field is inherently ambiguous. If analyzed locally, the interpreted direction of translation will be perpendicular to each edge as represented by the thin arrows. If analyzed globally, the interpretation will be a rigidly translating square as represented here by the thick arrows.

fering interpretations, the visual system defaults to the locally defined solution (Lorenceanu & Shiffrar, 1992; Shiffrar & Pavel, 1991).

If biological motion displays are analyzed globally, observers may be more likely to select global interpretations of multiple aperture displays when the global interpretation is consistent with human locomotion. We therefore set out to determine whether observers could combine biological motion signals across disconnected apertures. To create differing local and global motion interpretations, we created stick figure renditions of a point light walker and placed them behind a set of apertures positioned so that the walker's joints were always hidden from view. The apertures were invisible so that line terminators provided locally unambiguous motion signals (Shimojo et al., 1989). A local interpretation of this display would be the perception of line segments translating along the length of the apertures, as in the barber pole illusion. A global interpretation would be the perception of a walker.

EXPERIMENT 1 Objects Behind Apertures

Are naive observers able to recognize human gait by globally combining spatially disconnected motion signals, or do they default to a local solution and perceive independently moving line segments? The walking stick figure generates a complex pattern of motion signals. If observers recognize the walker behind apertures, will their recognition abilities generalize to the perception of other complex yet nonbiological objects?

Method

Subjects. Sixty Rutgers University undergraduates participated in this experiment for credit toward a class requirement. All subjects had normal or corrected-to-normal visual acuity and were naive with respect to the hypothesis under investigation. The subjects were tested individually.

Stimuli. All stimuli were displayed on a Macintosh 21-in. (40 × 30 cm) RGB monitor with a 1,152 × 870 pixel resolution. A Macintosh Quadra 950 was used to generate and control the stimuli. The temporal display rate was controlled by Macromedia Director Version 4.0 animation software. This apparatus was used in all of the experiments reported here.

The stimuli consisted of animated sequences of simplistic line drawings made with a white line having a width of 1.5' of visual angle. Each stimulus was shown for 6 sec at a rate of four frames per second. Stimulus construction began with the creation and modification of a computer-generated point light walker (Cutting, 1978) so that the final pattern of point motion was consistent with a man walking on a treadmill through two complete gait cycles. The walker subtended a 3.6° of visual angle (DVA) height and had a 1.6 DVA step size from the subjects' viewing distance of 80.5 cm. Once the point light walker was created, we connected the points on each frame with 10 line segments to construct a walking stick figure, as shown in Figure 2.

A car stimulus was created by importing a six-frame sequence of a moving car and eliminating everything from each frame except the outline of the car. This outline was then simplified and replaced with 15 white line segments and two white circles. The resulting car figure had a 3.6 DVA length and a 1.8 DVA height. The car translated over a distance of 0.85 DVA in 1.75 sec, creating a

translation velocity of 0.48 DVA/sec. The car translated back and forth across this distance twice.

A scissors stimulus was created by animating six-frame sequences consisting of a line drawing of a pair of scissors. The scissors consisted of four line segments and two ovals. The scissors' length was 3.6 DVA and their width varied from a minimum of 1.8 DVA when the scissors were closed to a maximum of 2.35 DVA when the scissors were fully open. The mouth of the scissors opened to a maximum angle of 25° with an angular velocity of 14.3°/sec. The scissors opened and closed twice within a trial. As with the walker, both the car and scissors stimuli were shown at a rate of four frames per second.

The aperture set, shown in Figure 2, was constructed from 14 rectangular windows that were each 0.14 DVA wide and 0.71 DVA long. Windows were positioned end to end at 90° angles to create three sawtooth patterns. The apertures had the same black color and luminance as the rest of the display and therefore were invisible. This aperture shape was selected because it kept the walker's joints continuously hidden and created a distinctive local motion interpretation. This aperture set was used in every condition, although its overall orientation was inverted in the car condition and rotated 90° in the scissors condition in order to hide all stimulus discontinuities (such as corners). When viewed through this set of apertures, the walker had six visible edges on any frame, the scissors had eight visible edges on any frame, and the car had an average of seven visible edges (varying between a minimum of six and a maximum of nine edges).

In the static conditions, one frame was selected from each of the three animated figure sequences. The selected frames had the largest figure width (1.6 DVA for the walker and 2.35 DVA for the scissors) and number of visible edges. Each single, static frame was shown for 6 sec on each trial. The selected frames are shown in Figure 2.

Pilot studies were used to ensure that all three figures, when unoccluded, were easily identifiable. Naive observers were shown each figure for 1 sec and were asked to identify the figure and rate their ease of identification. All observers accurately and effortlessly identified every figure.

Procedure. The subjects were seated in front of the display monitor and were told that they would see a brief display. They were asked to watch the entire display until it disappeared and then to describe what they had seen. In a 3 × 2 between-subjects design, each subject viewed the car, walker, or scissors in either the dynamic or the static condition. Ten subjects were randomly assigned to each condition. After receiving instructions, the subjects viewed a black screen with a fixation point in the center. Once the subject indicated that he/she was prepared, the experimenter initiated the trial by pressing a button on a mouse device. At that point, the fixation point disappeared and one of the six possible displays appeared, centered at the same location. After 6.0 sec, the aperture display disappeared. The subjects then verbally described what they had seen into a tape recorder. Two reasons motivated the use of a free identification procedure. First, this measure replicates that used by Johansson (1976) in his classic perceptual studies of biological motion. Second, in most previous tests of the global analysis of biological motion, subjects have known what form they needed to locate. We selected a task within which subjects would be unable to develop task-specific strategies or to use prior knowledge to direct their motion analyses.

Once all 60 subjects had completed the experiment, the tape-recorded responses were interpreted by the experimenter as well as by a second rater who was blind with respect to the conditions and goals of the study.

Results and Discussion

In each condition, the number of subjects who correctly identified the underlying form of each stick figure

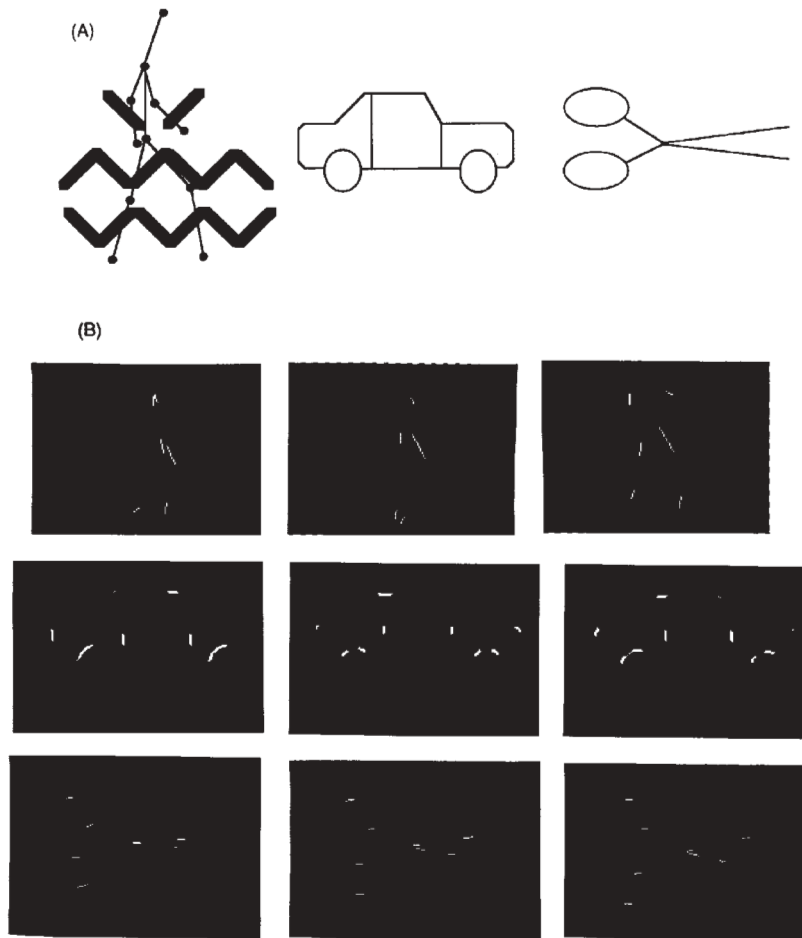


Figure 2. (A) The stick figure stimuli used in all experiments. The walker is shown with the apertures. (B) Three frames from each dynamic display. The top row shows the walker, the second row shows the car, and the third row shows the scissors. The last frame in each row shows the stimulus used in the static condition of Experiment 1.

was tabulated. Responses fell into two distinct categories. In nearly all cases, subjects either named an object or described a field of independently translating line segments. The response classifications made by the experimenter and rater were in complete agreement. The results, shown in Figure 3, are described as the percentage of subjects correctly identifying each stimulus. In the dynamic condition, all 10 subjects correctly identified the walker, 1 subject correctly identified the car, and no subject identified the scissors. Typical descriptions of the dynamic car and scissors stimuli included: "lines moving," "birds," "worm-like things that got longer," "undulating lines," and "a bunch of lines." Movement is critical in this recognition task, since none of the subjects correctly identified the walker, car, or scissors in the static condition. This result is consistent with previous studies of object recognition demonstrating that stationary objects with occluded corners cannot be identified (Biederman, 1987). The static condition results are also consistent with point light walker studies in which observers have significant difficulties in recognizing static point light walkers and

are unable to discriminate between static male and female walkers (Kozlowski & Cutting, 1977).

The results from the dynamic condition are consistent with the hypothesis that the visual system analyzes human gait globally. When the local and global interpretations of a dynamic, multiple aperture display differ, observers select the global interpretation when it is consistent with a human walker. The ceiling level of performance in the dynamic walker condition is impressive, since the walker's joints were not visible. In previous studies, researchers have examined perceptual speed and accuracy with point light walkers when the lights are positioned in between the walker's joints. The perception of off-joint walkers is less accurate than that of on-joint walkers, since observers of off-joint walkers cannot discriminate male from female walkers (Cutting, 1981) and are slower and less accurate in their perception of instrumental actions (Dittrich, 1993). Similarly, infants require more time to perceptually organize off-joint displays (Bertenthal et al., 1987).

Since it is difficult to compare recognition rates across differing objects and motions, the nonbiological figures

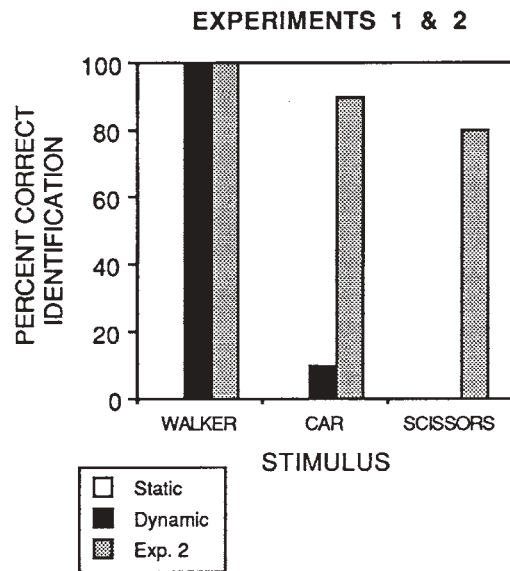


Figure 3. The results of Experiments 1 (black and white bars) and 2 (gray bars). All subjects were able to correctly identify the upright dynamic walker seen through either visible or invisible apertures. Although most subjects could recognize the dynamic car and scissors through visible apertures (Experiment 2), they could not do so through invisible apertures. Subjects were unable to recognize the static displays, and hence the white bars are not visible in the graph.

used in this study were highly familiar. The car stimulus was constructed so that it always had a greater number of visible lines than did the walker stimulus. Thus, the present results cannot be explained in terms of the number of visible lines. The scissors stimulus was used to determine whether nonrigid stimuli, in general, are analyzed more globally than rigid stimuli. Since observers were unable to recognize the scissors stimulus, the results do not support a generalized, global analysis of all nonrigid stimuli. Taken together, the present results support previous proposals that the ability to analyze forms globally may not be shared by all objects (Bertenthal, 1993; Farah, 1992; Shiffrar, 1994). Finally, these results support the interpretation of results from masked point light walker studies that global analyses are sufficient for the perception of biological motion (Bertenthal & Pinto, 1994).

EXPERIMENT 2 A Recognition Control

An alternative explanation of subjects' inability to recognize the partially occluded car and scissors in Experiment 1 is simply that these figures were not recognizable. The walker, car, and scissors represent different classes of objects that vary significantly along several dimensions. Were subjects more likely to recognize the walker display simply because it was a better representation of the real object? To control for the recognizability of the nonbiological stimuli, we conducted a short ex-

periment in which the local and global motion interpretations of these stimuli were compatible.

Occluded objects have two different types of surface boundaries: real boundaries and temporary boundaries created by the occluding surface. Accurate object recognition requires that the visual system rely on real boundaries and discount temporary boundaries (Kanizsa, 1979). One of the ways in which the visual system distinguishes real from temporary boundaries involves the use of occlusion cues. When depth cues suggest that contour terminators are the temporary result of another occluding surface, those terminators do not influence image interpretation (Shimojo et al., 1989). On the other hand, in the absence of compelling occlusion cues, terminators play a defining role in image interpretation. In Experiment 1, no depth cues were present, because the apertures were invisible. Under these conditions, the line segments appeared to translate along the length of the apertures, as in the barber pole illusion. In the present experiment, we added occlusion cues to the multiple aperture display so that the terminators would be correctly classified as temporary and subsequently discounted from the motion analysis. To this end, we simply increased the luminance of the area surrounding each aperture so that T junctions would be created where the lines intersected the now visible apertures. This manipulation should eliminate terminator-based interpretations of the moving lines and thereby the conflict between the local and global interpretations of the walker, car, and scissors. If the partially occluded car and scissors are recognizable stimuli, subjects should be able to accurately identify them under these visible aperture conditions.

Method

Subjects. Ten Rutgers University undergraduates participated in this experiment for credit toward a course requirement. All subjects were naive with respect to the hypothesis under investigation and none had participated in the previous experiment. All subjects had normal or corrected-to-normal visual acuity. The subjects were run individually.

Stimuli and Procedure. The stimuli from the dynamic condition of Experiment 1 were used in this experiment. The only modification to these stimuli was an increase in the luminance, from black to light gray, of the area surrounding the apertures. This created visible T junctions wherever the moving lines intersected the apertures. All other aspects of the stimuli remained unchanged.

According to a within-subjects design, every subject saw all three dynamic stimuli. The order of stimulus presentation was counterbalanced and randomized. To help them accurately segment the apertures from the figures, the subjects were told that they would see an image through a set of windows. As before, the subjects were instructed to view each 24-frame animated sequence and to describe what they had seen.

Results and Discussion

The results, shown in Figure 3, indicate that both the biological walker and the nonbiological car and scissors stimuli could be identified when viewed through visible apertures. The visible apertures created a situation in which the local and global motion interpretations of the display were consistent. Under these conditions, most

subjects could identify all of the moving figures. However, not every subject recognized both the car and the scissors stimuli. We attribute this less than perfect performance to our use of monocular rather than binocular occlusion cues. Binocular disparity has a significantly greater influence on terminator classification than monocular cues do (Shimojo et al., 1989). Nonetheless, the present results clearly suggest that the dynamic, partially occluded car and scissors are recognizable figures.

EXPERIMENT 3 Orientation Specificity

One of the defining characteristics of the perception of point light walkers is orientation specificity. Observers can readily identify a point light walker when the walker is in an upright, canonical orientation, but recognition rates drop significantly if a point light walker is shown upside down (Barclay, Cutting, & Kozlowski, 1978; Bertenthal, 1993; Dittrich, 1993; Pavlova, 1989; Sumi, 1984). For example, adult observers can detect an upright but not an upside-down point light walker presented in a mask of scrambled point limbs (Bertenthal & Pinto, 1994). If the present aperture-based walker displays are analyzed by the same visual mechanism(s) as are point light displays, recognition accuracy should drop significantly with upside-down aperture displays.

Method

Subjects. Ninety Rutgers University undergraduates participated in this experiment for credit toward a course requirement. All subjects were naive with respect to the hypothesis under investigation and none had participated in the previous experiments. All subjects had normal or corrected-to-normal visual acuity. The subjects were run individually.

Stimuli and Procedure. An opaque 30×30 DVA white screen with a 7-DVA-diameter circular opening in the center was used to

cover the monitor edges to minimize local framing effects. All stimuli were viewed through this circular window. The stimuli consisted of the same animated walker, car, or scissors figure used in the dynamic condition of Experiment 1. As before, all figures were viewed through the same set of invisible apertures. There were three possible stimulus orientations for each stimulus. The 0° orientation condition was a replication of the dynamic condition of Experiment 1, with the walker, car, and scissors in their canonical orientations. It should be noted that the "canonical" orientation for the scissors was subjectively selected. In the 90° and 180° conditions, all three figures were rotated 90° or 180° from canonical, respectively. The 90° and 180° orientations were created by simply rotating the rectangular display monitor. This manipulation ensured that the spatial resolution of the stimuli remained constant across changes in orientation.

In a 3×3 between-subjects design, 90 naive observers briefly viewed and tried to identify the dynamic walker, car, or scissors behind invisible apertures at one of three orientations. All other stimulus parameters were identical to those of Experiment 1.

Results and Discussion

The results, shown in Figure 4, indicate that recognition of the walker was strongly influenced by stimulus orientation. All subjects correctly identified the upright walker, but only 3 recognized the horizontally oriented walker and only 1 recognized the inverted walker. Incorrect responses in the 90° and 180° walker conditions included "intersecting lines," "birds flying," "two sets of lines making circular motion" and "little dotted lines." None of the subjects in the scissors conditions correctly identified that figure in any orientation. The car stimulus was correctly identified in its canonical orientation only once. As before, incorrect responses to the car and scissors stimuli involved various descriptions of independently moving line segments. These results are consistent with the orientation specificity of biological motion analyses. Moreover, these results refute an explanation of subjects' ability to recognize the walker on the basis of the presence of a hierarchical organization, in general,

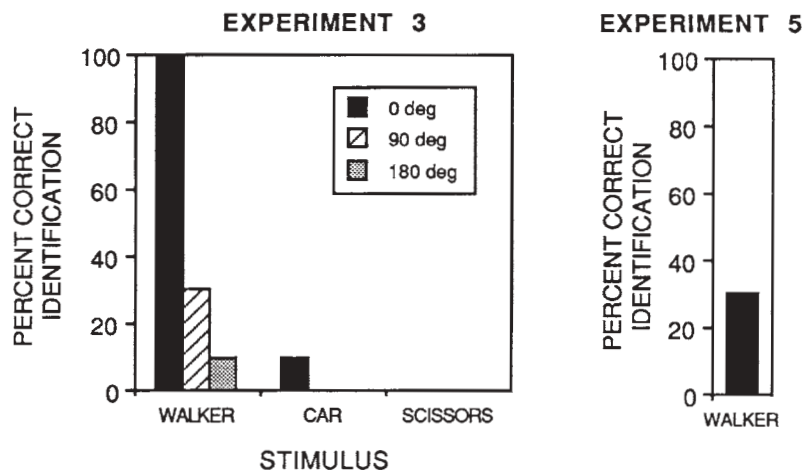


Figure 4. The results of Experiment 3 suggest that the recognition of the walker through apertures is orientation dependent. The results of Experiment 5 suggest that limb orientation cues may facilitate walker recognition.

rather than a human form, in specific. Finally, this pattern of results supports the hypothesis that the aperture-based walker display may be analyzed by the same mechanisms that analyze point light walker displays.

EXPERIMENT 4
Speed

The previous studies involved slowly moving objects, since the animation sequences were shown at a rate of four frames per second. Were the present results simply a function of the use of relatively long stimulus durations? Or would these effects generalize across a larger temporal window? To address this question, we varied the speed of the walker.

Method

Sixty Rutgers University undergraduates participated in this experiment for credit toward a course requirement. All subjects were naive with respect to the hypothesis under investigation, and none had participated in any of the previous experiments. All subjects had normal or corrected-to-normal visual acuity and were run individually.

The stimulus consisted of the dynamic walker viewed through invisible apertures, as in Experiment 1. However, unlike in the first experiment, there were six different stimulus durations. The shortest stimulus duration (SD) was 62 msec or 16 frames per second, and each subsequent SD was simply doubled, to yield 125, 250, 500, 1,000, and 2,000 msec. The 250-msec stimulus duration replicated the frame rate used in the previous experiments. Temporal limitations of our apparatus negated the reliable use of shorter stimulus durations. No interstimulus interval was used. According to a between-subjects design, each subject observed the walker display at one of six possible SDs. The total number of frames in each display remained fixed at 24.

Results and Discussion

The results, shown in Figure 5, indicate that stimulus duration influences walker recognition. Walker recognition was at or near ceiling when the stimulus duration was 125–500 msec. This interval corresponds to one complete step cycle in 1.5–6.0 sec and includes realistic walking speeds (Barclay et al., 1978). Performance dropped at rates above and below this temporal window. By way of informal comparison, we also showed a small group of subjects the car and scissors stimuli through invisible apertures. Irrespective of the speed at which these inanimate stimuli were shown, the subjects were always unable to identify them.

One interesting aspect of these walker results is that performance remained fairly good out to unusually long stimulus durations. There is reason to believe that the spatiotemporal window over which apparent motion can be seen is extended in the analysis of biological motion (DeSilva, 1926). However, the 40% correct identification at the longest SD suggested that observers were using more than motion information to identify the walkers. In the following experiment, we investigated whether form information would facilitate the recognition of walkers behind apertures.

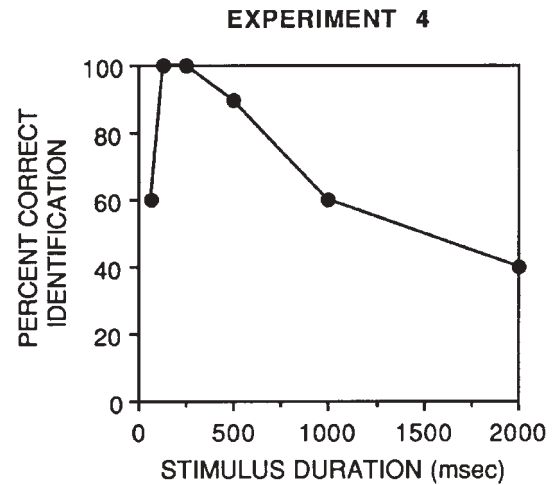


Figure 5. The results of Experiment 4. Stimulus duration influences walker recognition.

EXPERIMENT 5
Object Form

Traditionally, studies of the visual perception of biological motion have focused on the role of motion cues. However, recent behavioral and neurophysiological studies suggest that form also plays an important role in the perception of biological motion. Neurons in the STPa of the macaque exhibit an increased rate of firing when presented with displays consisting of precise combinations of biological forms and motions (Perrett et al., 1990). Under control conditions, these neurons do not show an increased firing rate if a biological movement is generated by a nonbiological form such as a metal bar. Moreover, apparent motion studies further support the importance of form cues in the analysis of biological motion. In classic demonstrations of apparent motion, two sequentially presented elements appear as a single moving element. This situation is ambiguous, because there is an infinite number of possible paths of motion connecting any two elements. Observers generally overcome this ambiguity by perceiving the shortest possible path of apparent motion, even if that path requires quite large variations in object form. However, when viewing photographs of a human model in different positions, observers perceive apparent motion paths that are consistent with the movement limitations of the human body rather than paths that traverse the shortest possible distance (Shiffrar & Freyd, 1990, 1993). Thus, biological forms can influence the interpretation of biological motions.

We wondered whether subjects' recognition of the walker at unusually long stimulus durations might reflect the influence of form. Moreover, although observers always perform well in tasks involving the detection of point light walkers in a mask, they generally do not perform at ceiling levels. The stick figure walkers, unlike

point light walkers, directly provide orientation cues for each walker limb. These orientation cues might facilitate the perceptual organization of the walkers. To test this hypothesis, we created a stimulus having a biological form but nonbiological motion.

Method

Ten Rutgers University undergraduates participated in this experiment for credit toward a course requirement. All subjects were naive with respect to the hypothesis under investigation, and none had participated in any of the previous experiments. All subjects had normal or corrected-to-normal visual acuity and were run individually.

The stimulus consisted of a single frame of the walker that was rigidly translated. This frame was identical to the one used in the static walker condition of Experiment 1 and was shown through the same stationary aperture set. However, in this experiment, the fixed walker was rigidly translated horizontally back and forth across a distance of 0.93 DVA at a velocity of 0.53 DVA/sec. As before, subjects viewed the display for 6 sec and then described what they had seen to the experimenter.

Results and Discussion

The results, shown in Figure 4, indicate that 30% of the observers correctly identified the rigidly translating stimulus as arising from a human form. Some observers reported that the figure appeared to be a strangely contorted man riding an invisible skateboard. This result suggests that the limb orientation cues provided by the stick figure stimulus did facilitate the accurate interpretation of the walker display, confirming the importance of global body form (Heptulla Chatterjee, Freyd, & Shiffrar, 1996). However, the fact that only a few of the observers recognized the walker suggests that form cues may play a smaller role than that played by motion cues in the visual analysis of biological motion. This interpretation is consistent with physiological studies suggesting that form plays a significant although secondary role in the response of STPa neurons to biological motion displays (Oram & Perrett, 1994).

GENERAL DISCUSSION

In a series of experiments, we investigated how the visual system interprets biological motion displays. Previous studies using masked point light walkers have generated conflicting conclusions regarding whether local or global analyses are sufficient for the visual analysis of biological motion (Bertenthal & Pinto, 1994; Mather, Radford, & West, 1992). We examined the global analysis of biological and nonbiological motions with a multiple aperture display in which the motion of every visible contour was equally ambiguous. Observers in our experiments easily recognized upright, walking stick figures viewed through apertures. This finding supports previous proposals that global analyses are sufficient for the perception of point light walkers within a mask (Bertenthal & Pinto, 1994). The present findings also suggest that human observers may be even more sensitive to biological motion than has previously been suggested. Johansson's original point light walker displays are ambiguous

because the dozen or so points can be grouped together in many different ways. The motion of each individual point is unambiguous. When a walking stick figure is viewed through apertures, every motion signal is ambiguous. Given this extra degree of ambiguity, the performance demonstrated by subjects in the present experiments suggests an impressive degree of sensitivity.

When observers viewed simplified line drawings of moving, nonbiological objects through invisible apertures, they were generally unable to recognize the underlying figures. This confirms previous findings that observers usually select local interpretations of moving images even when a global interpretation is consistent with a rigid object (Lorenceanu & Shiffrar, 1992; Shiffrar & Pavel, 1991). If subjects had been using an object rigidity constraint, or the bias to select image interpretations consistent with rigid objects (Ullman, 1979), then the dynamic car condition should have led to the best recognition performance. Thus, local analyses appear to dominate over global analyses and the use of an object rigidity constraint in the perception of moving, nonbiological objects.

Recognition rates differed greatly between the biological walker and nonbiological car and scissors displays. One possible interpretation of this difference is that observers have more familiarity with walkers than with cars or scissors and that this difference led to differences in recognition. Evidence against this interpretation comes from Experiment 2, in which observers were able to recognize the figures when they viewed them through visible apertures. Thus, if familiarity differences caused our results, familiarity must interact with motion perception in a relatively subtle way. An alternative interpretation of our results is that human observers analyze biological motion displays more globally than they analyze nonbiological displays (Bertenthal, 1993; Farah, 1992; Shiffrar, 1994; Vaina, Lemay, Bienfang, Choi, & Nakayama, 1990).

One of the main controversies in current visual research concerns the extent to which form and motion analyses interact. Psychophysical studies have suggested that form and motion are processed independently (Burt & Sperling, 1981; Krumhansl, 1984). Neurophysiological studies confirm this segregation with the proposal that spatial analyses and object recognition are conducted in separate cortical pathways (Ungerleider & Mishkin, 1982). If form and motion analyses proceed independently, why are observers able to recognize point light or aperture walkers? One possible reason is suggested by recent studies showing that although form and motion analyses may proceed independently in the ventral and dorsal pathways, respectively, these pathways converge in area STP (Baizer, Ungerleider, & Desimone, 1991). Perrett and his colleagues have identified cells in this area that are sensitive to biological motion (Oram & Perrett, 1994; Perrett et al., 1990). Thus, form and motion analyses may converge for the analysis of moving biological forms.

Another possible explanation is that form and motion analyses interact for the analysis of complex, nonrigid stimuli such as point light walkers but not for rigid stimuli (Cutting et al., 1988). The scissors display served as

a preliminary test of this hypothesis. Even though the scissors were nonrigid and fairly complex, observers were unable to recognize them through invisible apertures. This result suggests that the interaction of form and motion may be limited to biological displays in particular rather than nonrigid displays in general. On the other hand, while nonrigid, the scissor motion was not as complex as the walker motion. Observers may accurately perceive more complex, nonrigid, nonbiological motions. Additional research is needed to address these hypotheses.

REFERENCES

- ADELSON, E., & MOVSHON, J. A. (1982). Phenomenal coherence of moving visual patterns. *Nature*, **300**, 523-525.
- BAIZER, J., UNGERLEIDER, L., & DESIMONE, R. (1991). Organization of visual inputs to the inferior temporal and posterior parietal cortex in macaques. *Journal of Neuroscience*, **11**, 168-190.
- BARCLAY, C. D., CUTTING, J. E., & KOZLOWSKI, L. T. (1978). Temporal and spatial factors in gait perception that influence gender recognition. *Perception & Psychophysics*, **23**, 145-152.
- BERTENTHAL, B. I. (1993). Perception of biomechanical motions by infants: Intrinsic image and knowledge-based constraints. In C. Granrud (Ed.), *Carnegie Symposium on Cognition: Visual perception and cognition in infancy* (pp. 175-214). Hillsdale, NJ: Erlbaum.
- BERTENTHAL, B. I., & PINTO, J. (1994). Global processing of biological motions. *Psychological Science*, **5**, 221-225.
- BERTENTHAL, B. I., PROFFITT, D. R., KRAMER, S. J., & SPETNER, N. B. (1987). Infants' encoding of kinetic displays varying in relative coherence. *Developmental Psychology*, **23**, 171-178.
- BIEDERMAN, I. (1987). Recognition by components: A theory of human image understanding. *Psychological Review*, **94**, 115-147.
- BURT, P., & SPERLING, G. (1981). Time, distance, and feature trade-offs in visual apparent motion. *Psychological Review*, **88**, 171-195.
- CUTTING, J. E. (1978). A program to generate synthetic walkers as dynamic point-light displays. *Behavior Research Methods & Instrumentation*, **10**, 91-94.
- CUTTING, J. E. (1981). Coding theory adapted to gait perception. *Journal of Experimental Psychology: Human Perception & Performance*, **7**, 71-81.
- CUTTING, J. E., & KOZLOWSKI, L. T. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, **9**, 353-356.
- CUTTING, J. E., MOORE, C., & MORRISON, R. (1988). Masking the motions of human gait. *Perception & Psychophysics*, **44**, 339-347.
- DESILVA, H. R. (1926). An experimental investigation of the determinants of apparent visual motion. *Journal of Experimental Psychology*, **37**, 469-501.
- DITTRICH, W. H. (1993). Action categories and the perception of biological motion. *Perception*, **22**, 15-22.
- FARAH, M. (1992). Object recognition: You may mistake your wife for a hat, but not for a word. *Current Directions in Psychological Science*, **1**, 164-169.
- HEPTULLA CHATTERJEE, S., FREYD, J., & SHIFFRAR, M. (1996). Configurational processing in the perception of apparent biological motion. *Journal of Experimental Psychology: Human Perception & Performance*, **22**, 916-929.
- HOFFMAN, D. D., & FLINCHBAUGH, B. E. (1982). The interpretation of biological motion. *Biological Cybernetics*, **42**, 195-204.
- JOHANSSON, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, **14**, 201-211.
- JOHANSSON, G. (1976). Spatio-temporal differentiation and integration in visual motion perception. *Psychological Review*, **38**, 379-393.
- KANIZSA, G. (1979). *Organization in vision: Essays on gestalt perception*. New York: Praeger.
- KOZLOWSKI, L. T., & CUTTING, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics*, **21**, 575-580.
- KRUMHANS, C. (1984). Independent processing of visual form and motion. *Perception*, **13**, 535-546.
- LORENCEAU, J., & SHIFFRAR, M. (1992). The role of terminators in motion integration across contours. *Vision Research*, **32**, 263-273.
- MAREY, E.-J. (1972). *Movement* (E. Pritchard, Trans.). New York: Arno Press. (Original work published 1895)
- MATHER, G., RADFORD, K., & WEST, S. (1992). Low level visual processing of biological motion. *Proceedings of the Royal Society of London: Series B*, **249**, 149-155.
- MATHER, G., & WEST, S. (1993). Recognition of animal locomotion from dynamic point-light displays. *Perception*, **22**, 759-766.
- MINGOLLA, E., TODD, J., & NORMAN, J. F. (1992). The perception of globally coherent motion. *Vision Research*, **32**, 1015-1031.
- ORAM, M., & PERRETT, D. (1994). Responses of anterior superior temporal polysensory (STPa) neurons to "biological motion" stimuli. *Journal of Cognitive Neuroscience*, **6**, 99-116.
- PAVLOVA, M. (1989). The role of inversion in perception of biological motion pattern. *Perception*, **18**, 510.
- PERRET, D., HARRIES, M., MISTLIN, A. J., & CHITTY, A. J. (1990). Three stages in the classification of body movements by visual neurons. In H. B. Barlow, C. Blakemore, & M. Weston-Smith (Eds.), *Images and understanding* (pp. 94-107). Cambridge: Cambridge University Press.
- SHIFFRAR, M. (1994). When what meets where. *Current Directions in Psychological Science*, **3**, 96-100.
- SHIFFRAR, M., & FREYD, J. J. (1990). Apparent motion of the human body. *Psychological Science*, **1**, 257-264.
- SHIFFRAR, M., & FREYD, J. J. (1993). Timing and apparent motion path choice with human body photographs. *Psychological Science*, **4**, 379-384.
- SHIFFRAR, M., & PAVEL, M. (1991). Percepts of rigid motion within and across apertures. *Journal of Experimental Psychology: Human Perception & Performance*, **17**, 749-761.
- SHIMOJO, S., SILVERMAN, G., & NAKAYAMA, K. (1989). Occlusion and the solution to the aperture problem for motion. *Vision Research*, **29**, 619-626.
- SUMI, S. (1984). Upside-down presentation of the Johansson moving light-spot pattern. *Perception*, **13**, 283-286.
- ULLMAN, S. (1979). The interpretation of structure from motion. *Proceedings of the Royal Society of London: Series B*, **203**, 405-426.
- UNGERLEIDER, L., & MISHKIN, M. (1982). Two cortical visual systems. In D. Ingle, M. Goodale, & R. Mansfield (Eds.), *Analysis of visual behavior* (pp. 549-586). Cambridge, MA: MIT Press.
- VAINA, L., LEMAY, M., BIENFANG, D., CHOI, A., & NAKAYAMA, K. (1990). Intact "biological motion" and "structure from motion" perception in a patient with impaired motion mechanisms: A case study. *Visual Neuroscience*, **5**, 353-369.
- WALLACH, H. (1935). Über visuell wahrgenommene Bewegungsrichtung [On the visually perceived direction of motion]. *Psychologische Forschung*, **20**, 325-380.
- WALLACH, H. (1976). On perceived identity: I. The direction of motion of straight lines. In H. Wallach (Ed.), *On perception* (pp. 201-216). New York: New York Times Book Co., Quadrangle.
- WEBB, J. A., & AGGARWAL, J. K. (1982). Structure from motion of rigid and jointed objects. *Artificial Intelligence*, **19**, 107-130.
- WILSON, H., FERRERA, V., & YO, C. (1992). A psychophysically motivated model for two-dimensional motion perception. *Visual Neuroscience*, **9**, 79-97.

(Manuscript received March 22, 1995;
revision accepted for publication February 22, 1996.)