Structural processing in biological motion perception

Hongjing Lu
Departments of Psychology and Statistics, University of California, Los Angeles, USA

To investigate the basis for biological motion perception, structural and motion information were manipulated independently in a dynamic display using a novel stimulus with multiple apertures. Performance was compared in discrimination of global motion (translation and rotation) and biological motion. When structural information in the display was eliminated but motion information was intact, human observers were able to perceive global motion yet were at chance in discriminating walking direction of biological movement. In contrast, when the display provided even noisy and impoverished structural information, walking direction became identifiable. The present findings thus provide direct psychophysical evidence that motion information is insufficient and structural information is necessary for the identification of walking direction in biological movement. These findings imply that computational models must utilize a structural representation of the human body to account for perception of biological movements.

Keywords: biological motion, aperture problem, structural processing, motion processing


Introduction

Understanding the basis for the perception of biological movements (e.g., walking or running) is a fundamental issue in vision science. Many species have been found to be exceptional at recognizing motion patterns generated by other living organisms, presumably due to the ecological importance of biological motion for survival in a dynamic visual world (Blake, 1993; Cutting & Kozlowski, 1977; Fox & McDaniel, 1982; Regolene, Tommasi, & Vallorigara, 1999; Simion, Regolin, & Buolf, 2007). Determining how the visual system achieves this perceptual feat has challenged generations of researchers. Physiological studies have revealed that a distributed cortical network is involved in perceiving complex actions (Grossman & Blake, 2002; Nelissen, Vanduffel, & Orban, 2006; Puce & Perrett, 2003; Rizzolattti & Craighero, 2004), including early visual cortex in the occipital lobe, action-sensitive regions in the superior temporal sulcus, and the mirror neuron system in the frontal lobe.

Behavioral studies have shown that human abilities in recognizing biological motion are noteworthy not only for their speed and accuracy, but also for their flexibility and robustness, as humans can recognize biological motion from impoverished visual input. A particularly compelling stimulus is the point-light display, which depicts human actions using only discrete joints in a motion sequence (Johansson, 1973). Despite the lack of a detailed human body form and the virtual absence of point-light displays in natural scenes, observers vividly perceive complex actions (Dittrich, 1993; Dittrich, Trosclair, Lea, & Morgan, 1996) and accurately identify characteristics of an actor, such as identity (Cutting & Kozlowski, 1977; Troje, Westhoff, & Lavrov, 2005), emotional state (Dittrich et al., 1996; Roether, Omlor, Christensen, & Giese, 2009), and gender (Kozlowski & Cutting, 1977; Mather & Murdoch, 1994; Pollick, Kay, Heim, & Stringer, 2005; Troje, 2002). Moreover, perception of human action remains robust even when point lights are embedded in a noisy background (Bertenthal & Pinto, 1994; Cutting, Moore, & Morrison, 1988; Neri, Morrone, & Burr, 1998), assigned with random contrasts (Ahlstrom, Blake, & Ashlstrom, 1997), or associated with scrambled depth (Bulthoff, Bulthoff, & Sinha, 1998; Lu, Tjan, & Liu, 2006).

The remarkably rapid, accurate, and robust perception of biological movement has inspired a great deal of research addressing one fundamental question: What visual information is crucial for perceiving biological movement? Several studies have provided evidence that motion computation plays a central role in biological motion perception (Cutting et al., 1988; Mather, Radford, & West, 1992; Thurman & Grossman, 2008). For example, studies using point-light displays have demonstrated that impressive recognition performance can be achieved from dynamic but not static displays, suggesting that humans may identify biological motion solely using local motion signals from joints (Johansson, 1973; Mather et al., 1992). On the other hand, studies have also shown that performance in recognizing biological movement is impaired substantially when a point-light actor is displayed upside down (Bertenthal & Pinto, 1994; Pinto & Shiffrar, 1999; Sumi, 1984). Beintema and Lappe (2002) showed that humans can perceive biological movement even when point-light positions are jittered dynamically to eliminate local motion information. Evidence from analyses of classification images has revealed comparable
correlations between all the point lights and recognition performance, implying a global representation of biological movement (Lu & Liu, 2006). The latter class of findings provides support for the hypothesis that structural or form information is also important in the visual processing of biological movement. For the purposes of the present paper, structural information refers to form cues that explicitly encode human body structure in the visual stimuli. These cues, which are readily accessible in the displays of a human stick figure or of a full body shape based on the orientation of limbs, are distinct from the information retrieved from motion using the mechanisms of structure from motion.

Overall, empirical tests of visual information crucial to perception of biological movement have so far been inconclusive. For example, although the inversion effect has long been attributed to impaired global structural processing in biological motion perception, Troje and Westhoff (2006) reported a significant inversion effect even when point lights were spatially scrambled to disrupt structural information in the display. This finding suggested that two mechanisms, global form analysis and a local motion trajectory detector (focusing especially on the feet), contribute to the inversion effect in biological motion perception. Similarly, Chang and Troje (2009) compared coherent and spatially scrambled point-light displays in various tasks in order to characterize different properties of global structural and local motion processes in biological motion perception.

Although point-light displays have served as extremely useful research tools, they have limitations. It is important to note that point-light displays in fact provide structural information (albeit very impoverished) about human body structure via spatial grouping. For example, observers with prior knowledge are able to group point lights to form a human figure in a static display (Lu & Liu, 2006). Thus, when point-light displays are used as stimuli, it is difficult to determine whether or not biological movement can be identified from motion cues alone. Second, point-light displays tend to promote feature-tracking mechanisms for motion analysis, making it difficult to use such stimuli to test contributions from other potential motion mechanisms. Finally, the rarity of point-light displays in natural scenes calls into question the generalizability of experimental findings based on these laboratory stimuli.

As Shiffrar, Lichtey, and Heptulla Chatterjee (1997, p. 51) have pointed out, “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.” Shiffrar et al. introduced a novel display in which visual stimuli were generated from a stick figure of a walker viewed through a set of fence-like rectangle occluders. The researchers found that human observers were able to identify a walker, but not a car or scissors, through invisible occluders. A key characteristic of Shiffrar et al.’s stimulus display is its inherent ambiguity in local motion signals, typically referred to as the aperture problem. The component of the line motion orthogonal to the line’s orientation can be clearly perceived; however, the component of motion parallel to the line’s orientation is not observable. One way to overcome the ambiguity in motion estimation is to integrate motion information over space and thereby perceive global motion patterns. For example, Amano, Edwards, Badcock, and Nishida (2009) and Mingolla, Todd, and Norman (1992) used a large number of circular apertures to investigate how the human visual system integrates a set of local motion measures to perceive a globally coherent motion pattern, such as translation. An appealing advantage of multiple aperture stimuli is that each aperture contains one or more lines/contours for which orientations and motion velocities can be independently specified.

The present study generalized the multiple aperture display in order to independently manipulate motion and structure information in biological movement stimuli. Figure 1 illustrates an example in natural scenes, which the multiple aperture display aims to mimic in the laboratory. In this example, the visual system needs to effectively integrate local motion signals to perceive multiple complex motion patterns, including the movement of the walker and the moving background respectively. In the multiple aperture display, the target that appeared behind a large number of circular apertures was a stick figure of a human actor, similar to that used in the study by Shiffrar et al. (1997). Orientation of each line in an individual aperture can be determined using the structure of the human body appearing behind the apertures. The local motion of each line in the window can be generated by the corresponding velocity of biological movement. In this paper, five experiments are reported to investigate the roles of structural and motion information in biological movement perception.

**Experiment 1: A comparison between global motion and biological motion**

Experiment 1 investigated human ability to spatially integrate local motion signals using two basic types of stimuli: global motion stimuli (specifically, translation and circular motion) and biological motion stimuli. Are naïve observers able to recognize each of these three motion patterns (translation, circular and biological) by pooling local motion signals across the space? If so, will humans exhibit superior recognition performance for biological motion relative to global motion, given the well-established robustness of human biological motion perception?
Method

Subjects

Eight University of California, Los Angeles (UCLA) undergraduate students participated in the experiment for course credit. All observers were naïve concerning the hypothesis under investigation and had normal or corrected-to-normal visual acuity.

Apparatus

Stimuli were presented on a Viewsonic monitor with a refresh rate of 75 Hz and resolution of 1024 × 768 pixels. At the viewing distance of 57 cm (maintained via a chin rest), each pixel subtended 1.98 min-arc. The monitor was calibrated with a Minolta CS-100 photometer. A lookup table was constructed to allow linear division of a luminance range, 0–146.5 cd/m², into 256 programmable intensity levels. Experiments were conducted in a dim room. Poser 4 software (MetaCreation Inc.) was used to create the walker stimulus; Matlab and PsychToolbox (Brainard, 1997; Pelli, 1997) were used to present the stimuli.

Stimuli

As shown in Figure 2, the stimuli consisted of 729 drifting Gabor elements located in a 27 × 27 grid, with separation of 0.4° within a square window subtending 11.6°. Each Gabor element (0.4° by 0.4°) was composed of an oriented sinusoidal grating of 5.1 cycles/deg frequency, windowed by a stationary Gaussian function with standard deviation of 0.08 degree. The contrast of Gabor elements was 0.4. Orientations of Gabor elements were randomly assigned in Experiment 1. Drifting speed and direction of each Gabor were determined by a sine frequency, windowed by a stationary Gaussian function with standard deviation of 0.08 degree. The contrast of Gabor elements was 0.4. Orientations of Gabor elements were randomly assigned in Experiment 1. Drifting speed and direction of each Gabor were determined by a sine
function of the angle between the Gabor orientation and its assigned motion velocity. The stimuli consisted of 40 image frames presented at a rate of 53 ms/frame.

The Gabor elements were categorized into two groups: foreground elements and background elements. Foreground elements lay on the trajectory of a human walker, whereas background elements lay off the trajectory. The walker moved as if on a treadmill for one walking cycle (including two steps) at the center of the display window. Three-dimensional coordinates were known for thirteen joints: head, left/right shoulder, left/right elbow, left/right hand, left/right thigh, left/right knee, and left/right foot for each frame; thus, joint movements in the stimulus were well defined. The walker was generated by connecting joints according to the human body hierarchy with a line width of 0.8. The size of the walker was 5.5 horizontally by 8.9 vertically. For each frame, center locations of foreground elements overlapped with the walker; the remaining element locations without overlap were defined as background elements.

In Experiment 1, each Gabor element, regardless of whether it was foreground or background, had randomly assigned orientations within a range of [0°, 180°]. However, different motion flows were assigned to foreground elements and background elements, respectively. As shown by the red arrows in Figure 3, motion velocities for background elements were assigned randomly in each frame, with a fixed motion speed of 0.8/sec. In contrast, as shown by blue arrows in Figure 3, motion velocity of foreground elements followed one of three global motion patterns: translation (top column), circular (middle column),
or biological (bottom column). In the conditions of translation and rotation, motion speeds of foreground elements were the same as background elements, 0.8°/sec. Motion directions of foreground elements were determined by global motion pattern (i.e., left/right translation, or else clockwise/counter-clockwise rotation). For the biological motion condition, foreground motion velocities were computed via spatial interpolation of the known joint motion velocities for each frame.

Procedure

The task was motion direction discrimination: left/right for translation, clockwise/counter-clockwise for rotation, and leftward/rightward for walking direction in biological motion. Prior to testing, observers were presented with three practice blocks to allow familiarization with the walking stimulus. In the first block, only foreground elements of walking stimuli were presented for 8 trials (4 trials walking leftward and 4 trials walking rightward). Observers were asked to judge the walking direction. In the second block, foreground elements and 50% background elements were presented for 12 trials of the direction discrimination task. In the third block, foreground elements and 100% background elements were presented with 20 trials for each of the three motion patterns (translation, rotation and walking). Feedback was provided for all practice trials.

Six blocks were included in the testing phase. Each motion condition was presented for two blocks of trials, with 50 trials in each block. The first three blocks included one block for each of the three conditions; their presentation order was randomly assigned for observers. For each observer, the final three blocks had the condition orders reversed relative to the first three blocks to eliminate any order effect. Observers were informed of the type of motion pattern (translation, circular, or biological motion) before each block started.

On each trial, the motion stimulus was presented after displaying the first frame for 0.53 sec. The location of the walker used to determine foreground/background elements was jittered around the center horizontally in a range of ±2.2° in each trial. A red fixation disk with a diameter of 0.36° was displayed at the center. After the stimulus disappeared from the screen, observers were asked to press one of the two response buttons to indicate perceived motion direction. The next trial started 1 sec after the observer made a response. No feedback was provided in the testing phase.

Results and discussion

In Experiment 1, Gabor orientations were randomly assigned, eliminating structural information in the stimuli for all three motion conditions. The results showed that observers were able to use global analysis of motion integration to succeed in the translation and circular conditions but completely failed to perceive biological movement in the walking discrimination task. Figure 4 depicts the percentage of responses correctly identifying motion direction in the three conditions. Observers achieved high levels of discrimination performance for translation and circular motion (79% and 93%, respectively) but were virtually at chance for biological motion (54%, no significant difference from chance performance, \( t(7) = 1.76, p = 0.12 \)).

The excellent discrimination performance observed in the translation and circular conditions is impressive. These results are consistent with many studies of motion perception using random dot kinematograms (Barlow & Tripathy, 1997; Burr, Morrone, & Vaina, 1998; Morrone, Burr, & Vaina, 1995) and multiple grating stimuli (Amano et al., 2009; Mingolla et al., 1992), which have shown that the human visual system is able to integrate motion information across space and orientation to perceive globally coherent motion patterns. The superior performance observed for circular compared to translation motion \( (t(7) = 3.34, p = 0.01) \), depicted in Figure 4, is consistent with previous findings in the literature on motion perception (Freeman & Harris, 1992; Lee & Lu, 2010) showing that humans exhibit greater motion sensitivity to rotation than translation.

In view of the strong performance for global motion, the catastrophic failure of discrimination performance for the biological motion condition is especially surprising. Performance on the biological motion task not only failed to exceed that on the global motion tasks but in fact fell to chance. This failure in perceiving biological motion certainly does not support the common assumption that

![Figure 4](http://jov.arvojournals.org/) The results of Experiment 1. Percent correct in identifying motion direction for each of three types of motion patterns. Error bars indicate standard error of the mean (SEM).
the human visual system is exceptional at recognizing biological motion from impoverished visual inputs. The chance performance for the biological motion condition indicates that because of the random assignment of orientations for foreground elements, the stimuli used in Experiment 1 in fact eliminated the most critical information for achieving successful recognition of biological motion—the structure of the human body.

On the face of it, the results of Experiment 1 might seem to contradict findings from the study by Shiffrar et al. (1997), summarized in the Introduction section. These investigators found that observers could readily identify partially occluded walkers, but not a car or scissors, through multiple apertures. However, the difference between the results from the two studies may be explained by considering whether or not structural information was included in the visual stimuli. It appears that when structural information is completely eliminated by randomly assigning orientations for elements, as was done in Experiment 1, human observers are unable to identify biological movement. However, if the stimulus maintains structural information of the human body, as was the case for the stick figure used in Shiffrar et al.’s study, observers may in fact be able to perceive biological movement successfully. Experiment 2 was designed to test this hypothesis using psychometric measures.

**Experiment 2: Biological motion perception with structural noise**

If structural information is a necessary condition for successful biological motion perception, we would expect to obtain chance performance in the absence of structural information, and a monotonic improvement in identification accuracy with an increase in the amount of structural information included in the visual stimuli. Experiment 2 introduced structural information by manipulating the orientations of foreground elements. The amount of structural information was controlled by the variation of these orientations.

**Method**

Eight fresh naïve observers (UCLA undergraduate students) participated in Experiment 2 for course credit. The stimuli from the biological motion condition of Experiment 1 were used. The only modification to these stimuli was that orientations of foreground elements were randomly assigned from a uniform distribution centered at the body structure orientation with a range of ±a. The range of this uniform distribution controlled the deviation of assigned orientation from intact body structure orientations. As shown in Figure 5A, if the range is 0°, orientations of foreground elements align perfectly with the body structure, and a human figure can be readily detected even in one static frame. If the range is ±40°, orientations of foreground elements are more distributed, and it is difficult to detect a human figure from one static frame.
frame (Figure 5B). If the range is ±90°, orientations of foreground elements are randomly assigned, generating the same condition as that used in Experiment 1.

For each foreground element, orientation deviation from body structure was determined by random sampling but was fixed over frames. Orientations of background elements were randomly assigned within a range of [0°, 180°], also fixed over frames. The assignment of motion information was the same as in Experiment 1. The assigned motion for foreground elements was determined by motion flow of the walker.

The basic task was to discriminate the walking direction. The general procedure, including practice and testing phase, was similar to the biological motion condition of Experiment 1. Within each session, five orientation ranges for foreground elements were tested, ±10°, ±20°, ±40°, ±60°, and ±90°, in a blocked design. Each condition was tested in two blocks with 30 trials per block. The order of conditions for the first versus second series of five blocks was reversed for each individual observer. Observers were asked to identify walking direction without feedback. The walker’s location was jittered in a range of ±2° horizontally from trial to trial.

Results and discussion

The results were shown as red bars in Figure 6. When structural information was completely eliminated by enlarging the orientation range to ±90°, accuracy fell to chance level, replicating the comparable finding in Experiment 1 (mean accuracy 54%, t(7) = 2.05, p = 0.08). However, when structural information was included even with large uncertainty (orientation range ±60°), identification accuracy was imperfect but still well above chance (mean accuracy was 72%, t(5) = 6.95, p < .01). Furthermore, with reduced uncertainty in structural information, observers could readily identify biological movement when the orientation range for foreground elements was ±40°, ±20°, and ±10°, with the performance level near ceiling (mean accuracy 92%, 98%, and 99%, respectively). Using the condition of orientation range ±10° as the reference group, pairwise comparisons showed significantly decreased discrimination performance when orientation range was ±60° and ±90° (p < 0.01) but similar performance when the orientation range was ±20° (t(7) = 0.55, p = 0.6) and ±40° (t(7) = 1.65, p = 0.14).

These results demonstrate that the amount of structural information included in the visual stimuli plays a critical role in determining discrimination performance for biological motion perception. The results of Experiment 2 provide a coherent explanation of the apparent discrepancy between the results obtained in Experiment 1 and the findings reported by Shiffrar et al. (1997). In Experiment 1, biological motion perception was impossible because structural information was eliminated by randomizing orientations, but stick walkers were identifiable in the Shiffrar et al.’s study because no structural noise was added to their stimulus. The present findings are also consistent with the results of an additional experiment reported by Shiffrar et al. (Experiment 5), which revealed that limb orientation cues facilitate accurate interpretation of the walker display using the stick figure stimulus.

Experiment 3: Biological movement perception with random motion information

In Experiment 2, structural information about the human body was directly manipulated via orientations of foreground elements, but motion information was left intact by assigning proper motion velocities to foreground elements according to the motion flow of a walker. In this type of situation, human observers could rely on Gestalt grouping rules, such as colinearity, to connect foreground elements in order to form a global percept of a moving human body, even when a large amount of structural noise is introduced into the stimulus. Experiment 3 was designed to investigate whether human observers could still discriminate walking direction when local motion information was randomly assigned and to assess how discrimination performance varies as a function of structural noise when motion information is randomly assigned in the display.
Method

Twelve fresh naïve UCLA undergraduate students participated in Experiment 3 for course credit. The experimental design was identical to that of Experiment 2, except that motion velocities were randomly assigned to each foreground element. Motion directions for both foreground and background elements were assigned randomly in each frame, with a fixed speed of 0.8°/sec. To manipulate structural information, we introduced orientation noise in five orientation ranges for foreground elements, ±10°, ±20°, ±40°, ±60°, and ±90°, the same values used in Experiment 2.

Results and discussion

Discrimination performance as a function of orientation range is shown as green bars in Figure 6. When structural information was presented with a low amount of noise (i.e., the orientation range was ±10° or ±20°), human observers could readily identify biological movement even though the motion information provided by foreground elements was random (mean accuracy 97% and 92%, respectively). However, as the level of noise obscuring body structure increased further, the absence of motion information in foreground elements reduced human recognition performance significantly. When the orientation range was ±40°, although observers still achieved above-chance performance when random motion was assigned (mean accuracy 66%), their accuracy was much lower than the comparable condition in Experiment 2 with intact motion assignments (mean accuracy 92%). This difference was reliable ($t(18) = 4.62, p < 0.01$). This finding suggests that, when structural information is noisy or difficult to access in the stimulus display, motion plays an important role in identifying walking direction of biological movement.

A similar result was obtained when the orientation range was ±60°, in which case the absence of motion information reduced observer performance to 57%. Although this value was above chance ($t(11) = 3.2, p < 0.01$), it was significantly lower than the 72% accuracy observed in the comparable condition of Experiment 2 with the motion information available ($t(18) = 4.25, p < 0.01$). As shown in the last bar in Figure 6, when both structural and motion information were eliminated in the display (i.e., orientation range of ±90° with random motion assignments), observers’ performance in discriminating walking direction was reduced to chance level (52%, $t(11) = 1.04, p = 0.32$).

These findings suggest that a global analysis of biological motion using structural information can to some extent overcome the randomness in the local motion level. When structural information was presented with a low amount of noise, we found that discrimination performance was not affected by local motion information. This result is consistent with other studies demonstrating that the ability to identify biological movement is not disturbed by variations in low level visual information, such as point lights assigned with random contrasts (Ahlstrom et al., 1997), or point lights associated with scrambled depth (Bulthoff et al., 1998; Lu et al., 2006). However, when noise creates a large amount of uncertainty in the structural analysis, local motion plays an important role in cooperating with the noisy structural information to aid in perceiving biological motion, as illustrated in the conditions of ±40° and ±60°. These findings of Experiment 3 also demonstrate the limit of the capability to perform biological motion analysis solely using structural information. When the structural analysis cannot be carried out with certainty, human observers are able to integrate it with motion analysis to achieve more accurate performance in recognizing biological movements.

Figure 6. The results of Experiments 2 and 3, showing percent correct as a function of orientation range of foreground elements. Red bars depict results from Experiment 2 when the assigned motion for foreground elements was determined by motion flow of a walker. Green bars depict results from Experiment 3 when the assigned motion for foreground elements was random from frame to frame.

Experiment 4: Biological movement perception with high signal-to-noise ratio

All the above three experiments illustrate the crucial role of structural information in determining whether the visual system is able to discriminate walking direction in biological movements. It is possible, however, that the
chance level performance in Experiment 1 might be given an alternative explanation in terms of sheer task difficulty. Task difficulty can be manipulated by employing various signal-to-noise ratios (i.e., proportion of the foreground elements that are walker patches versus background elements that are noise). Tolerance to this type of noise may be much lower when perceiving biological movements rather than global motion. Experiment 4 aimed to test this possibility by increasing the ratio of foreground walker elements to background noise elements. If task difficulty plays the most important role in affecting observer’s performance, we would expect to observe performance levels well above chance with high signal-to-noise ratio, despite the lack of structural information in the visual stimuli.

Method

Twenty-four fresh naïve UCLA undergraduate students participated in Experiment 4 for course credit. The stimuli from the biological motion condition of Experiment 1 were used. The only modification to these stimuli was to reduce the number of background elements to the minimum. As shown in the top panel of Figure 7, stimuli only included elements located on the trajectory of a human in the whole walking cycle, consisting of 161 drifting Gabor elements. Compared to 729 elements in the stimuli used in Experiment 1, this setup increased the ratio of signal foreground element versus noise background elements by a factor of 4.5.

Three conditions were included in Experiment 4, varying assignments of orientations and velocities to foreground elements: random orientations and random motion assignments (termed “random orientation and random motion condition”); random orientations and veridical walking motion (termed “random orientation and biological motion condition”); and orientations randomly assigned from a uniform distribution centered at the body structure with a range of ±60° and walking motion for foreground elements (termed “noisy orientation and biological motion condition”). Background elements were generated in the same way as in Experiment 1.

Observers were asked to identify walking direction without feedback. A blocked design was used to measure observer performance for the three conditions, as in Experiment 2. The general procedure, including practice and testing phase, was the same as in Experiment 2.

Results

Figure 7 depicts discrimination performance for the three experimental conditions. Despite increasing the ratio of signal foreground elements versus noise background elements by a factor of 4.5, human observers were still unable to identify walking direction in the absence of structural information. When orientations of foreground elements were randomized, the mean accuracy was 0.50 with veridical biological motion, comparable to the performance with random motion (random orientation and biological motion vs. random orientation and random motion, $t(23) = 0.64, p = .53$). This result indicates that the
chance performance observed in Experiment 1 was primarily due to the lack of structural information, not the low signal-to-noise ratio in the stimuli. However, if a small amount of structural information was included in the stimuli by setting orientations of foreground elements within a range of ±60° of human body structure, discrimination accuracy increased to 0.64 despite the noisiness of the structural information (t(23) = 3.95, p < .01), replicating the findings of Experiment 2.

**Experiment 5: Biological movement perception with perturbed phase information**

In the multiple aperture stimuli used in previous experiments, the motion of the limbs occurred at the veridical position in the spatiotemporal domain. Limb position was thus always canonical, which could provide possible cues to retrieve human form information using the mechanism of structure from motion. To test this possibility, Experiment 5 measured human discrimination performance with a walker stimulus in which the relative phase between the different joins was randomly offset (i.e., phase information of limbs was scrambled). If the structure-from-motion explanation were correct, we would expect that observers would show poorer discrimination performance for phase-scrambled walker stimuli than for the phase-intact walker stimuli.

**Method**

Seventeen fresh naïve UCLA undergraduate students participated in Experiment 5 for course credit. The stimuli and general procedure were similar to Experiment 4. In the phase-scrambled condition, the walker was generated by connecting joints undergoing the same motion trajectory, but with their temporal phases scrambled. One phase-scrambled walker without background elements (see the movie in Figure 8) was shown for the purpose of demonstration. In the phase-intact condition, the walker was generated in the same way as in previous experiments. Two levels of structural noise were included in the experiment: random orientations with the noise range of ±90°, and orientations with the range of ±60° around the connected limbs of the walker.

**Results**

As shown in Figure 8, although observers performed better for the phase-intact walker than for the phase-scrambled walker, the differences were very small. When the orientation range was ±60°, there was no significant difference between the two conditions (intact: .66; phase-scrambled: .62, t(16) = 1.1, p = 0.14). The above-chance performance in the phase-scrambled condition agrees with the finding that observers can retrieve information about direction from scrambled point-light displays (Troje & Westhoff, 2006). However, when structural information was made noisy by adding a large orientation noise, the small difference observed between phase-intact and phase-scrambled conditions indicates that the motion of the limbs occurring at the veridical spatiotemporal position was not sufficient to evoke a structure-from-motion mechanism that would enable discrimination of walking directions in biological movements. When orientations were assigned randomly (i.e., the orientation range was ±90°), observer discrimination performance was at chance level for both phase-intact and phase-scrambled conditions (intact: .50, phase-scrambled: .49).

**General discussion**

In a series of experiments, structural and motion information were manipulated independently in a moving sequence using multiple aperture displays. The results demonstrated that the accessibility of structural information is a necessary condition for discriminating walking direction of biological movements. Experiment 1 compared discrimination performance for global motion (translational and circular) with biological motion. When
structural information in the display was eliminated but motion information was intact, human observers were able to readily perceive translational and circular motion yet were at chance in recognizing biological movement. However, when the display provided even noisy structural information, biological motion could be identified quite accurately, as shown in Experiment 2. These results reveal that unlike perception of global motion, perception of biological movement is not a direct consequence of spatiotemporal integration of local motion information. Rather, the perceived structure of the human body provides a fundamental constraint that enables integration of motion information in this visual task.

Experiment 3 demonstrated that, to a certain extent, changing local motion information from biological motion to random motion did not affect identification performance for biological movement perception. This result indicates that as long as biological motion perception can be achieved through global analysis using structural information, the visual system is immune to random variations in low-level visual information, including local motion cues. However, this ability to rely on structural information has a limit. If a large amount of structural noise is embedded in the stimulus, thereby increasing the uncertainty of the structural analysis, then the visual system combines motion analysis with structural analysis to achieve robust recognition performance for biological movements.

Experiment 4 ruled out the possibility that a low proportion of signal foreground elements versus noise background elements could explain the chance performance for biological motion observed in Experiment 1. Rather, the lack of structural information appears to be the essential reason for the failure of discriminating walking direction in biological movements. Experiment 5 examined whether temporal perturbation of the underlying walker could affect discrimination performance when structural information was noisy. The small performance difference observed between the phase-intact and phase-scrambled condition indicated that the retrieval of form cues using structure-from-motion mechanisms may not play an important role in discriminating walking direction using the multiple aperture stimulus. Instead, structural information may guide the use of structure-from-motion mechanisms to integrate motion information for each limb and to segment motion information for different limbs. If structure information is completely eliminated from the stimuli, the visual system is not able to effectively elicit structure-from-motion mechanisms to perceive biological motion.

The present study thus provides direct psychophysical evidence that motion information is insufficient, and structural information is necessary, for the identification of biological motion. This finding is consistent with neuropsychological evidence. Vaina, Lemay, Bienfang, Choi, and Nakayama (1990) found that a patient with bilateral lesions involving the posterior visual pathways showed very poor performance for early motion tasks (e.g., coherent motion perception, speed discrimination) yet showed intact biological motion perception. Conversely, Cowey and Vaina (2000) found that a hemianopic patient with a lesion in ventral extrastriate cortex was unable to identify biological motion despite the fact that her motion perception was normal.

It appears that the fundamental difference between point-light displays and multiple-aperture displays is the source of ambiguity embedded in the visual stimulus. For classical point-light displays (without masking), ambiguity is largely associated with structural information because there are many different ways to group point lights to form different body structures. In contrast, the motion of each individual point light is not ambiguous because the correspondence problem is relatively easy to solve in such a short-range motion stimulus.

By contrast, multiple-aperture displays introduce ambiguity inherent in the local motion information for each element. Ambiguity in structural information can be controlled by the experimenter using orientation assignment. It follows that resolution of one of the main controversies in the area—the relative importance of motion and structural analysis in biological movement perception—may depend on the source of ambiguity in the stimuli. For point-light displays with at most a small number of masking dots, motion analysis plays a relatively important role because of the considerable ambiguity in structural grouping, as shown in a number of studies (Cutting et al., 1988; Mather et al., 1992; Thurman & Grossman, 2008). For point-light displays with a large number of masking dots, additional motion ambiguity arises due to the correspondence problem, in addition to structural ambiguity. Furthermore, the contribution from structural analysis varies depending on specific masking paradigms used in experiments (Bertenthal & Pinto, 1994; Pinto & Shiffrar, 1999) because the type of mask determines the information available in the stimulus. For example, Bertenthal and Pinto (1994) used masking dots that shared the same motion trajectories as walker point lights but scrambled spatial locations to rule out the possibility that the human form could be detected solely from the movements of individual point lights. For multiple-aperture stimuli, structural analysis makes a relatively greater contribution to perception of biological motion due to the aperture problem, which cannot be solved by local motion analysis. The latter type of stimuli thus provides an opportunity to investigate the minimum requirements for motion and structural analysis in biological movement perception.

From a computational perspective, the present findings support the hypothesis that the representation of biological movement cannot be solely based upon the motion trajectory of joints. Evidence consistent with this conclusion has also been provided by a study reported by Hunt and Halper (2008), suggesting that computational models of action perception must utilize structural
information. However, the format of the structural representation employed for biological motion perception remains unclear. Some current computational models assume a series of posture templates as the structural representation (Giese & Poggio, 2003; Lange & Lappe, 2006). An alternative possibility is that the structural representation takes the form of a structural description, in which human limbs are represented as arguments of predicates corresponding to spatial relationships between these parts. Future research should focus on identifying the format of structural representations that can flexibly construct meaningful groupings and thus guide motion integration and pattern analysis in biological movement perception.

**Acknowledgments**

This research was supported by a grant from the National Science Foundation (NSF BCS-0843880). I thank Alan Lee and Thach Nguyen for their help in data collection.

Commercial relationship: none.

Corresponding author: Hongjing Lu.

Email: hongjing@ucla.edu.

Address: Department of Psychology, University of California, Los Angeles, 405 Hilgard Ave., Los Angeles, CA 90095-1563, USA.

**References**


