THE VISUAL PERCEPTION OF HUMAN LOCOMOTION

Ian M. Thornton
University of Oregon, Eugene, OR, USA

Jeannine Pinto
Rutgers University, Newark, NJ, USA

Maggie Shiffrar
Rutgers University, Newark, NJ, USA and UMR CNRS, Université de la Méditerranée, Marseille, France

To function adeptly within our environment, we must perceive and interpret the movements of others. What mechanisms underlie our exquisite visual sensitivity to human movement? To address this question, a set of psychophysical studies was conducted to ascertain the temporal characteristics of the visual perception of human locomotion. Subjects viewed a computer-generated point-light walker presented within a mask under conditions of apparent motion. The temporal delay between the display frames as well as the motion characteristics of the mask were varied. With sufficiently long trial durations, performance in a direction discrimination task remained fairly constant across inter-stimulus interval (ISI) when the walker was presented within a random motion mask but increased with ISI when the mask motion duplicated the motion of the walker. This pattern of results suggests that both low-level and high-level visual analyses are involved in the visual perception of human locomotion. These findings are discussed in relation to recent neurophysiological data suggesting that the visual perception of human movement may involve a functional linkage between the visual and motor systems.

INTRODUCTION

Any animal's survival depends upon its ability to identify the movements of both prey and predators. As social animals, humans behave largely in accordance with their interpretations and predictions of the actions of others. If the visual system has evolved so as to be maximally sensitive to those factors upon which an animal's survival depends (Shepard,

1984), then one would expect to find that human observers are particularly sensitive to human movement. Several decades of perceptual research support this prediction. In a classic study of the visual perception of human movement, Johansson demonstrated that human observers can readily recognise extremely simplified deceptions of human locomotion (e.g. Johansson, 1973, 1975; Johansson, von Hofsten, & Jansson, 1980). Extending a technique first devised by Marey (1972) in 1895, Johansson created "point-light walker" displays by filming human actors with small light sources attached to their major joints. By adjusting the lighting, the resultant film showed only a dozen or so moving points of light, as shown in Fig. 1. Nevertheless, observers of these films report a clear and compelling perception of the precise actions performed by the point-light defined actors. Importantly, observers rarely recognise the human form in static displays of these films (Johansson, 1973). Subsequent research has demonstrated that our perception of the human form in such displays is rapid (Johansson, 1976), orientation specific (Bertenthal & Pinto, 1994; Pavlova, 1989; Sumi, 1984), tolerates random contrast variations (Ahlström, Blake, & Ahlström, 1997), and extends to the perception of complex actions (Dittrich, 1993), social dispositions (MacArthur & Baron, 1983), gender (Kozlowski & Cutting, 1977, 1978), and sign language (Poizner, Bellugi, & Lutes-Driscoll, 1981).

What neural mechanisms underlie the visual perception of human movement? Recent neurophysiological research suggests that
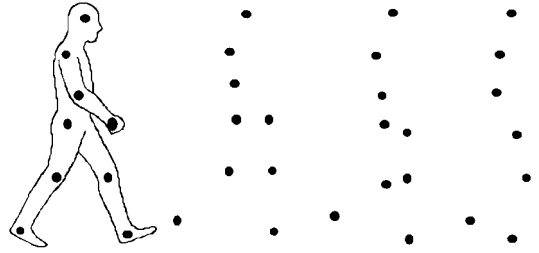

Fig. 1. Four static views of a point-light walker. The outline of the human body, shown in the first frame, is never shown in experimental stimuli. When presented statically, these displays are difficult to interpret. However, when set in motion, observers easily organise the complex patterns of point motion into a coherent perception of human locomotion.

relatively high-level integrative mechanisms may play a fundamental role in the visual analysis of human movement. For example, the superior temporal polysensory area (STP) of the macaque monkey, which receives input from both dorsal and ventral visual pathways (Baizer, Ungerleider, & Desimone, 1991), contains cells that appear to be selectively attuned to precise combinations of primate forms and movements (Perrett, Harries, Mistlin, & Chitty, 1990). Neurons in this area have also been shown to respond to Johansson point-light walker displays (Oram & Perrett, 1994). Furthermore, case studies of patients with extrastriate lesions sparing the temporal lobe demonstrate that individuals can lose their ability to perceive simple motion displays while retaining the perception of point light walker displays (Vaina, Lemay, Bienfang, Choi, & Nakayama, 1990; McLeod, Dittrich, Perrett, & Zihl, 1996).

A behavioural signature of high-level visual processes is their dependence upon global display characteristics. More specifically, most models of the visual system are hierarchical in nature (e.g. Van Essen & DeYoe, 1995; Zeki, 1993). Visual analyses at the lower levels of this hierarchy are thought to occur within brief temporal intervals and small spatial neighbourhoods. The results of these low-level or "local" analyses are then passed onto and processed by higher-level or more "global" mechanisms, which process information across larger spatiotemporal extents. Although local and global are difficult to define as absolute terms, most studies of the visual perception of human movement have defined local analyses as the computations conducted on individual points (joints) or point pairs (limbs). Global analyses are conducted over larger areas and generally involve half to an entire point-light walker. In the temporal domain, local motion processes are thought to be restricted to a window of 50msec or less (Baker & Braddick, 1985), while global motion processes may operate over much longer intervals.

Several psychophysical studies support the hypothesis that the visual perception of human movement depends upon a spatially global mechanism (e.g. Ahlström et al., 1997; Cutting, Moore, & Morrison, 1988). One approach to this issue involves masked point-light displays. In this paradigm, observers view displays containing a point-light walker that is masked by the addition of superimposed moving point lights. This mask can be constructed from multiple point-light walkers that are positionally scrambled so that the spatial location of each point is randomised. The size, luminance, and velocity of the points remain unchanged. Thus, the motion of each point in the mask is identical to the motion of one of the points defining the walker. As a result, only the spatially global configuration of the points distinguishes the walker from the mask. The fact that subjects are able to detect the presence as well as the direction of an upright point-light walker "hidden" within such a scrambled walker mask implies that the mechanism underlying the perception of human movement operates over large spatial scales (Bertenthal & Pinto, 1994). The spatially global analysis of human movement is further supported by studies of the aperture problem. Whenever a moving line is viewed through a relatively small window or aperture, its motion 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 (Wallach, 1935). The visual system can overcome this measurement ambiguity or aperture problem through local motion analyses (restricted to small spatial regions) or global motion analyses (that link information across disconnected spatial regions). When viewing a walking stick figure through a multiple aperture display, observers readily perceive global human movement. Under identical conditions, however, observers default to local interpretations of moving nonbiological objects and upside-down walkers (Shiffrar, Lichtey, & Heptulla-Chatterjee, 1997). This pattern of results suggests that the visual analysis of hu-

man locomotion can extend over a larger or more global spatial area than the visual analysis of other, nonbiological motions.

While the mechanism underlying the visual perception of human locomotion appears to conduct global analyses over space, its temporal characteristics remain unclear. Psychophysical researchers commonly use the phenomenon of apparent motion to investigate the temporal nature of motion processes. In classic demonstrations of apparent motion, two spatially separated objects are sequentially presented within a certain temporal range so that they give rise to the perception of a single moving object. Early studies demonstrated that apparent motion percepts depend critically upon the temporal separation of the displays (Korte, 1915; Wertheimer, 1912). When displays are separated by relatively long inter-stimulus intervals (ISIs), long-range apparent motion processes are thought to integrate information across the displays and to facilitate the perception of motion. On the other hand, when the frames in an apparent motion display are separated by brief temporal intervals (short ISIs), short-range processes are thought to underlie motion percepts (Anstis, 1980; Baker & Braddick, 1985). Long-range processes alone may conserve global cues to image structure such as object orientation (e.g. McBeath & Shepard, 1989), spatial frequency (e.g. Green, 1986), and perceptual grouping principles (e.g. Pantle & Petersik, 1980). Although there has been much debate concerning the precise nature of apparent motion phenomena (Cavanagh, 1991; Cavanagh & Mather, 1989; Petersik, 1989, 1991), the tradi-tional distinction between long- and short-range processes will be adopted here as it provides a useful framework within which to discuss temporal manipulations involving a single class of stimuli.

The perception of human movement in apparent motion displays provides an intriguing demonstration of the difference between short-range (temporally brief) and long-range (temporally extended) motion processes. In all apparent motion displays, the figure(s) shown in each display frame can be connected by an infinite number of possible paths. Under most conditions, however, observers typically report seeing only the shortest possible path of motion (e.g. Burt & Sperling, 1981). Yet, when humans move, their limbs tend to follow curved rather than straight trajectories. Given the visual system's shortest-path bias, will observers of human movement be more likely to perceive apparent motion paths that are consistent with the movement limitations of the human body or paths that traverse the shortest possible distance? This hypothesis has been tested previously with stimuli consisting of photographs of a human model in different positions created so that the biomechanically possible paths of motion conflicted with the shortest possible paths (Shiffrar & Freyd, 1990, 1993). For example, one stimulus consisted of two photographs in which the first displayed a standing woman with her right arm positioned on the right side of her head while the second photograph showed this same arm positioned on the left side of the woman's head. The shortest path connecting these two arm positions would involve the arm moving

through the head whereas a biomechanically plausible path would entail the arm moving around the head. When subjects viewed such stimuli, their perceived paths of motion changed with the Stimulus Onset Asynchrony (SOA) or the amount time between the onset of one photograph and the onset of the next photograph. At short SOAs, subjects reported seeing the shortest, physically impossible, motion path. However, with increasing SOAs, observers were increasingly likely to see apparent motion paths consistent with normal human movement (Shiffrar & Freyd, 1990). Conversely, when viewing photographs of inanimate control objects, subjects consistently perceived the same shortest path of apparent motion across increases in SOA. Importantly, when viewing photographs of a human model positioned so that the shortest movement path was a biomechanically plausible path, observers always reported seeing this shortest path (Shiffrar & Freyd, 1993). Thus, subjects do not simply report the perception of longer paths with longer presentation times. Moreover, observers can perceive apparent motion of nonbiological objects in a manner similar to apparent motion of human bodies. However, these objects must contain a global hierarchy of orientation and position cues resembling the entire human form before subjects perceive human-like paths (Heptulla-Chatterjee, Freyd, & Shiffrar, 1996). This pattern of results suggests that human movement is analysed by long-range motion processes that operate over large temporal intervals.

However, this conclusion appears inconsistent with the results of another series of apparent motion experiments (Mather, Radford, & West, 1992). These intriguing studies involved the presentation of synthesised point-light displays depicting the sagittal view of a person walking within a mask of randomly moving point lights. In some of these studies, observers reported whether the animated walker faced leftward or rightward in the picture plane. To create conditions appropriate for both long-range and short-range apparent motion, blank frames were added between the frames containing the masked walker. When the time between successive point-light walker frames (ISI) reached or surpassed 48msec, observers were unable to discriminate the two directions of walker motion. Since subjects could only perform the motion discrimination task under short-range apparent motion conditions, their perception of human movement appears to have depended upon local motion analyses. This finding suggests that the mechanism underlying the visual perception of biological motion analyses information within small temporal windows.

Thus, it is not yet clear whether the visual perception of human locomotion involves temporally local or global processes. Because the temporal studies cited differ significantly in methodology, their apparently conflicting results can not be unambiguously interpreted. Did the difference in results arise from methodological differences in display form, subject task, masking, or display duration? The goal of the following experiments was to resolve this interpretation limitation, and thereby to provide a better understanding of the mechanism underlying this perceptual behaviour. These

studies were motivated by the following assumption. If the neural mechanism subserving the visual perception of human locomotion operates over extended temporal windows, then subjects should be able to perform perceptual judgements of human locomotion under long-range apparent motion conditions.

EXPERIMENT 1: TRIAL DURATION

Why were subjects in the experiments of Mather et al. (1992) unable to determine a point-light walker's direction of motion under long-range apparent motion conditions? One possible reason concerns overall display duration. Johansson (1976) found that naive observers could identify a human form and its action from a point-light walker displayed for 200msec. However, the correct identification of a point-light walker presented within a mask requires longer display durations. Specifically performance in a direction discrimination task can fall to chance levels when masked point-light walkers are presented for less than 800msec (Cutting et al., 1988). In the experiments of Mather and his colleagues, the masked point-light walker was visible for as little as 240msec per trial. On the other hand, in the studies by Shiffrar and her colleagues (Heptulla-Chatterjee et al., 1996; Shiffrar & Freyd, 1990, 1993; Shiffrar et al., 1997), human movement displays were usually presented for several seconds. Thus, one possible explanation is that the use of brief display durations may lead to an underestimation of observers' perceptual capacities to interpret human movement. To examine this possibility, a modified replication of one of the studies conducted by Mather et al. (1992) was undertaken. Briefly, subjects performed a two-alternative forced-choice task in which they discriminated between rightward and leftward facing point-light walkers presented within a mask. The experimental modification involved the use of both long-duration and short-duration trials. If poor performance results from the use of excessively brief display durations, then performance in the long-duration trials should be superior to performance in the short-duration trials. Secondly, if above chance levels of performance are found, then the results of this experiment can be used to test whether low-level or high-level motion analyses are involved in the perception of human movement. More specifically, if performance at all ISIs is mediated exclusively by short-range motion processes, then performance should fall to chance levels with ISIs that extend beyond the temporal window for short-range analyses; namely, ISIs greater than approximately 50msec. If, however, the perception of human locomotion involves temporally extended motion analyses, then performance should remain well above chance with increases in ISI.

Method

Subjects

Three experienced psychophysical observers participated in this experiment. All observers had normal or corrected-to-normal vision. One subject was an author whereas the remaining

subjects were naive with regard to the purpose of this study.

Apparatus

All stimuli were displayed on a Macintosh 21" (40 × 30cm) RGB monitor with a refresh rate of 75Hz and a 1152 × 870 pixel resolution. Monitor output was controlled by a Macintosh Quadra 950. A chin rest was used to fix the subjects' viewing distance at 90cm from the monitor. The stimuli were presented in a 6.3° by 6.3° window positioned in the centre of the monitor. This window size closely replicated that used by Mather et al. (1992). This apparatus was used in both of the experiments reported here.

Stimuli

The stimuli were generated by modifying, in Think C version 7.0, a classic point-light walker algorithm (Cutting, 1978) together with a simultaneously presented mask of randomly moving dots (Cutting et al., 1988). Each animation frame consisted of 77 identical black dots displayed against a uniform, middle grey background. Eleven of these dots defined the walker while the remaining 66 dots defined the mask. Every dot, whether it belonged to the mask or the walker, was a 5 × 5 pixel square that subtended 6.1 min arc.

The simulated walker was displayed in profile as shown in Fig. 2. The dots that defined the walker were positioned on the simulated head, near shoulder, both elbows, both wrists, near hip, both knees, and both ankles of the walker (Cutting, 1978). As in previous masked point-light walker studies, the walker was always displayed with all 11 dots. That is, dots did not disappear when they would normally be occluded by the walker's torso or limbs. The removal of this natural occlusion cue minimised non-motion related cues to the location of the walker in the mask (Bertenthal & Pinto, 1994; Cutting et al., 1988; Mather et al., 1992). The mask dots themselves were placed randomly around the walker on a frame-by-frame basis. As a result, the dots defining the walker and the mask could only be distinguished from each other by their motion. Mather et al. (1992) nicely described these stimuli, when set in motion, as resembling a "figure striding through a light snowstorm".

The walker figure subtended 4.6° in height (head to ankle) and 2.4° in width at the most extended point of the step cycle. A complete stride cycle (i.e. the sequence of movements that occurs between two consecutive repetitions of a body configuration) was achieved in 40 animation frames. The duration of each frame was fixed at 40msec. As a result, when these frames were presented in immediate succession, a walking speed of 38 strides per minute was simulated. This speed falls within the range of 30–70 strides per minute associated with human walking under normal conditions (Inman, Ralston, & Todd, 1981). The walker figure did not translation across the screen but rather appeared to walk in place as if on a treadmill. On half of the trials, the walker faced and walked to the right while on the other half of the trials, the walker faced and walked to the left. The horizontal and vertical
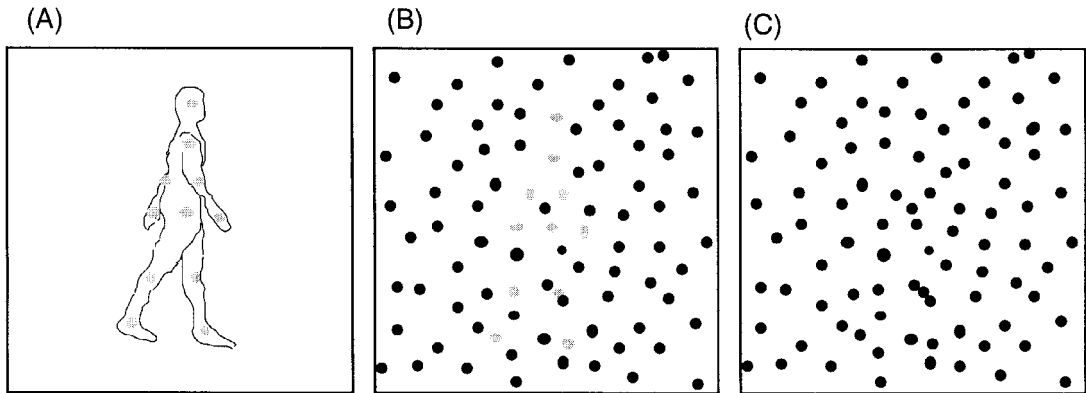
(A)   (B)   (C)

Fig. 2. The creation of a masked point-light walker display. Frame A illustrates a walker with 11 grey points fixed to each of the major body joints and the head. Frame B displays the grey point-light walker within a mask of black points. In the experimental stimuli, the walker points and mask points are identical, as shown in Frame C. The walker can be located within dynamic but not static displays.

position of the walker was randomised within the central display area on a trial-by-trial basis. The walker's position was constrained by the need to ensure that none of the dots defining the walker approached or exceeded the boundary of the display area. The starting position within a stride cycle (e.g. legs far apart or close together) was also randomised on each trial. These display manipulations ensured that subjects would not be able to identify the walker configuration simply by its presentation at a particular location or during a specific animation frame.

To manipulate the ISI, and thereby create long-range and short-range apparent motion, a blank frame was inserted between each of the animation frames. This blank frame contained no dots and was the same uniform grey as the background in the animation frames. Across trials, the duration of these blank frames was varied from 0msec (no blank frame) to 120msec in 15msec increments. This yielded a total of nine different Inter-stimulus Intervals (ISIs) of 0, 15, 30, 45, 60, 75, 90, 105 and 120msec.

There were two types of trials. A short-duration trial consisted of 20 animation frames and corresponded to half of a walker's stride cycle. This short-duration trial condition was selected in order to replicate the findings of Mather et al. (1992). Long-duration trials consisted of an 80-frame sequence and allowed for the presentation of two complete strides. Within each trial duration, the full range of ISIs was used. Trial duration was always equal to or greater than 800msec. More precisely, the overall duration of the 20 frame trials was 800msec when the ISI equalled 0msec and 3.2 sec when the ISI equalled 120msec. The 80 frame trials had durations as brief as 3.2sec and as long as 12.8sec when the ISI was 0 or 120msec, respectively.

Procedure

Subjects were seated in front of the display monitor and were told that they would see a point-light walker within a mask. They were instructed to determine, on each trial, if the walker's direction was to the left or right and then to press one of two buttons on a computer keyboard to indicate their decision. Responses could only be recorded after an animation sequence was completed. Subjects initiated the next trial by pressing another button on the keyboard. No feedback was provided during the practice or experimental sessions.

According to a within-subjects design, each subject completed four blocks of short-duration trials and four blocks of long-duration trials. These eight blocks were intermixed and their order was counterbalanced across subjects. Each block contained 10 trials at 9 different ISIs for a total of 90 trials. On average, subjects completed 1 block of trials in approximately 15 minutes. The order of the trials within each block was randomised independently for each subject. All subjects completed 18 practice trials before beginning each new block of experimental trials.

Results

The results, shown in Fig. 3, are plotted as the mean percentage of trials during which subjects correctly reported the walker's direction at each ISI level in both the short (20 frame) and long (80 frame) trial duration conditions. A 2 (Condition) × 9 (ISI) repeated measures ANOVA was used to analyse these data. A significant main effect of Condition [$F(1,2) =$
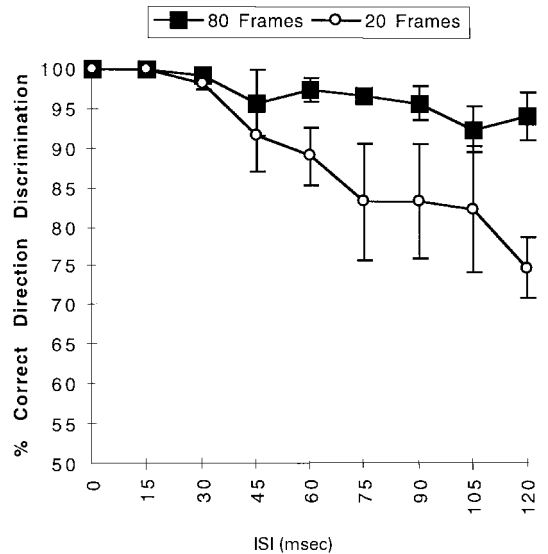

Fig. 3. The results of Experiment 1. The results are collapsed across subjects. Performance in the long-duration trial condition, indicated by the filled squares, remains high across variations in the ISI. Performance in the short-duration trial condition, shown by the empty circles, decreases with increasing temporal delays. The error bars represent the standard error of the mean.

23.07, $MSE = 33.8$, $P < .05$], was identified, with responses to 80 frame trials being more accurate ($M = 96.85$, $SD = 4.01$) than responses to 20 frame trials ($M = 89.26$, $SD = 11.16$). While there was also a significant main effect of ISI [$F(8,16) = 4.7$, $MSE = 40.5$, $P < .01$], this effect should be interpreted in the light of a Condition × ISI interaction [$F(8,16) = 3.12$, $MSE = 22.2$, $P < .05$]. To explore this interaction further, post hoc contrasts were used to compare Condition means at each level of ISI. This analysis revealed a significant divergence in performance by 60 msec [$F(1,16) = 4.7$, MSE

= 22.2, P < .05] with the short-duration trials remaining significantly below the long-duration trials for all ISIs beyond this point. Separate repeated measures ANOVAs confirmed this pattern of results with a strong main effect of ISI for the 20-frame condition [$F(8,16)$ = 4.37, MSE = 54.75, P < .01], but only a marginal effect for the 80-frame condition [$F(8,16$ = 2.54, MSE = 8.0, P < .064]. Finally, it is important to note that even the poorest performance, which occurred in the 20-frame condition when the ISI equalled the 120msec, was still significantly above chance [$t(2)$ = 6.55, P < .01].

Discussion

The results of this experiment clearly demonstrate that observers can perceive human locomotion under both long-range and short-range apparent motion conditions. More precisely, in the 20-frame condition, ceiling levels of performance were recorded when the temporal delay or ISI between the frames displaying the masked point-light walker was less than 60msec. This value is consistent with the 0 to 50msec temporal window associated with short-range apparent motion processes (Baker & Braddick, 1985). Beyond this point, performance dropped with increasing ISIs. This pattern of results replicates those of Mather et al. (1992, Expt. 2) in which direction discrimination performance dropped with ISIs greater than 48msec. However, in the present experiment, performance in the 80-frame trial duration condition remained relatively flat across increases in ISI. Since the long duration trial condition was constructed by simply increas-

ing the number of walker frames from 20 to 80, the responses of low-level motion detectors should have remained unchanged. Nonetheless, subjects were better able to determine the point-light walker's direction of motion under long-range apparent motion conditions when trial durations were extended beyond those used by Mather et al.

Although the pattern of results from the short trial duration condition is very similar to the pattern reported in Mather et al. (1992), absolute performance differs. Subjects in the current experiment performed the direction discrimination task more accurately than subjects in the direction discrimination experiment of Mather et al. This difference may reflect our use of only trained psychophysical observers. However, we have since replicated this same pattern of results with more than 20 naive observers (Pinto, Thorton, & Shiffrar, 1998). Superior overall performance may have also resulted from differences in frame duration. Each walker frame was displayed for 40msec in the current experiment but for only 24msec in the direction discrimination experiment by Mather and his colleagues. Thus, superior performance with longer frame durations is completely consistent with the hypothesis that subjects perform relatively poor perceptual judgements of masked human locomotion when displays are presented only briefly (Cutting et al., 1988).

Previous investigators of the visual perception of biological motion have used masked point-light walker displays to examine the spatial nature of this perceptual process. The results of their studies suggest that the percep-

tion of human movement involves spatially global analyses (Bertenthal & Pinto, 1994; Cutting et al., 1988). In earlier studies of the temporal characteristics of biological motion perception, researchers have varied the delay between photographs of a human model in different positions. The results of these studies support the existence of a temporally global mechanism (Shiffrar & Freyd, 1990, 1993). The current methodology involved a combination of these strategies, since a temporal delay was inserted between frames depicting a masked point-light walker. The current results therefore suggest that subjects can make subtle perceptual judgements about human locomotion even when these judgements require visual analyses that are global across both space and time. This findings is consistent with the hypothesis that a high-level mechanism, rather than low-level motion processes alone, underlies the visual perception of human movement.

However, it is important to note that the results of this experiment can not be convincingly interpreted as exclusively representing a high-level mechanism. That is, if performance in the long trial duration condition were solely the function of a temporally global analysis, then performance should have been independent of ISI. Yet, performance varied with ISI. One possible interpretation of this result is that local motion analyses may be involved in the perception of human movement. The goal of the following experiment was to determine more precisely whether low level motion analyses play a role in the visual perception of human movement.

EXPERIMENT 2: MASK COMPLEXITY

The mask used in Experiment 1 and in Mather et al. (1992) consisted of randomly moving points. Thus, the position of each point in the mask was uncorrelated from frame to frame. Since the walker points had pendular trajectories that simulated normal human locomotion, the position of these points was correlated across frames. As a result, the motions of the individual points of the mask and walker differed. These local differences were therefore available to low-level motion detectors and may have contributed to the detection of the walker in the mask. Therefore, a different type of mask is needed to eliminate the utility of low-level motion processes.

Previous research has shown that subjects can accurately discriminate the direction of a point-light walker in a mask even when the motion of each mask point mimics the motion of a walker point (Bertenthal & Pinto, 1994). These so-called "scrambled walker" masks are constructed by duplicating a point-light walker several times and then scrambling the starting position, but not the motion trajectory, of each point. This process yields a mask which might, for example, consist of points corresponding to seven left wrists plus seven right wrists plus seven left ankles plus seven heads, etc., and each having a randomly determined location within the 2D plane of the mask. Only the configuration of points that define the walker can be used to distinguish the walker from the mask. Thus, such "scrambled walker" masks more thoroughly camouflage human location than "random dot" masks (Cutting et

al., 1988). In other words, "scrambled walker" masks can be used to eliminate or drastically reduce the influence of low-level motion processes in the perception of point-light walkers.

If the visual analysis of human locomotion is global across both space and time, then subjects should be able to interpret a point-light walker within a scrambled walker mask even under conditions of long-range apparent motion. To test this prediction, subjects performed a modified replication of Experiment 1 in which the same point-light walker was presented within a scrambled walker mask rather than a random dot mask.

Method

The same three psychophysical observers from Experiment 1 served as subjects in this experiment. As before, two of the subjects were naive to the hypothesis under investigation.

The subjects' task in this experiment was identical to that of the previous experiment. The displays were also identical except for the motion trajectories of the dots making up the mask. In the previous experiment, the mask dots moved randomly. In this experiment, each dot in the mask had a motion trajectory that was identical to the trajectory of one of the dots defining the walker. This "scrambled walker" mask was created by generating six copies of the walker within the display area. The initial vertical and horizontal positions of each dot were then randomised within the display window. As a result, each mask dot had the same velocity as one of the walker dots but bore no predictable spatial relationship to any

other dot. As before, the mask dots also had the same size, colour, and luminance as the walker dots. The experimental procedure replicated that of Experiment 1.

Results

The results, shown in Fig. 4 as the mean percentage of trials during which subjects correctly reported the walker's direction at each ISI level, were analysed in a 2 (Condition) × 9 (ISI) repeated measures ANOVA. This yielded a significant main effect of Condition [$F(1,2) = 51.12$, $MSE = 32.1$, $P < .05$], with responses to 80-frame trials being more accurate ($M = 74.72$, $SD = 13.36$) than responses to 20-frame trials
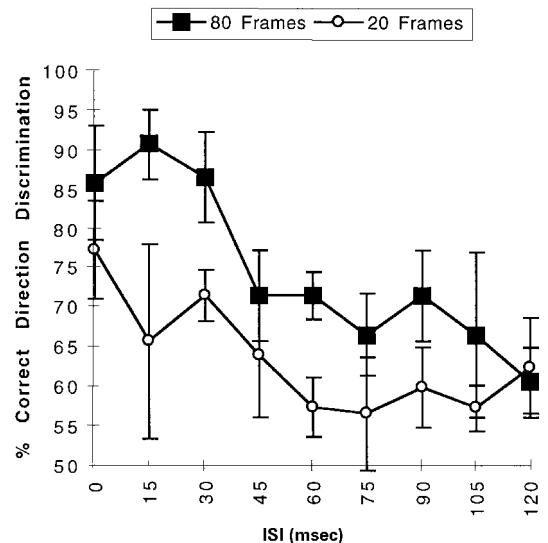

Fig. 4. The results of Experiment 2 collapsed across subjects. Performance in the long-duration trial condition (filled squares) is above chance for ISIs less than 90 msec and superior to performance in the short-duration trial condition (empty circles). Error bars represent the standard error of the mean.

($M = 63.7, SD = 11.82$). Unlike in Experiment 1, there was no Condition × ISI interaction. Separate analysis of the data from the two conditions revealed only a marginal main effect of ISI for the 20-frame condition [$F(8,16) = 2.46$, $MSE = 61.01$, $P < .06$] and a significant main effect of ISI for the 80-frame condition [$F(8,16) = 6.13$, $MSE = 53.65$, $P < .01$]. Polynomial contrasts revealed that this main effect had a strong linear component [$F(1,16) = 40.13$, $MSE = 53.65$, $P < .001$], reflecting a gradual drop in performance between the 0msec ($M = 85.83$, $SD = 12.58$) and 120msec ($M = 60.83$, $SD = 7.2$) ISI increments. T-tests indicated that performance in the 20-frame condition remained at chance levels (50%) for all ISI increments except 0 and 30msec. In contrast, in the 80-frame condition, performance remained significantly above chance (all ps $< .05$) for all ISIs except those of 105msec ($P = .13$) and 120msec ($P = .06$).

Discussion

Three general conclusions are suggested by the results of this study. First, performance in this direction discrimination task is better at long (80-frame) trial durations than at short (20-frame) trial durations. This finding further supports the hypothesis that poor performance in this task can stem from the use of trials presented over insufficient durations. Second, performance in the long-duration trial condition suggests that subjects can integrate motion correctly over large spatial and temporal extents in the analysis of human locomotion even when masking renders local motion sig-

nals uninformative. This finding clearly suggests that high-level or temporally global motion analyses are involved in the visual perception of human movement. Finally, comparison with the results of Experiment 1 demonstrates that the perception of a point-light walker is more difficult when it is presented within a mask of identically moving points than in a mask of randomly moving points. Local differences in motion trajectories are available in random dot masks but not in scrambled walker masks. These local motion differences may account for the performance differences between Experiments 1 and 2. This interpretation is further supported by the results of the long-duration trial condition in this experiment. Although performance was generally above chance, it also dropped with increasing ISI. The influence of low-level motion detectors is thought to decrease as temporal delays increase (e.g. Baker & Braddick, 1985). If so, when considered together, these results suggest that both low-level (Mather et al., 1992) and high-level (Bertenthal & Pinto, 1994; Shiffrar & Freyd, 1990, 1993) visual mechanisms may be involved in the visual perception of human locomotion

GENERAL DISCUSSION

The goal of this behavioural research project was to develop a better understanding of the mechanisms underlying the visual interpretation of human movement by examining the temporal characteristics of locomotion perception. In two experiments, subjects viewed

Johansson-like point-light walkers presented within a mask of moving points and reported the walker's direction of motion. Apparent motion displays were created by inserting blank frames of variable duration (or ISIs) between the walker frames. In Experiment 1, subjects viewed point-light walkers within a mask of randomly moving points over short and long trial durations. When only 20 walker frames were presented, performance dropped with ISIs greater than 60msec. This performance pattern replicates earlier findings (Mather et al., 1992). When the same masked walker was shown for 80 frames per trial, near-ceiling levels of performance were found across variations in ISI. This finding, that longer trial durations can improve performance, supports previous demonstrations that subjects report the perception of human movement under long-range apparent motion conditions (Heptulla-Chatterjee et al., 1996; Shiffrar & Freyd, 1990, 1993). When considered together, the results of this experiment suggest that the perceptual processes tapped by point light walker displays can operate over extended spatio-temporal neighbourhoods. Such global behaviour is generally considered to be a signature of mechanisms resisting within relatively late stages of the visual system.

In Experiment 2, the point-light walker was presented within a "scrambled walker" mask rather than in a "random dot" mask. As a result, the motion trajectories of the points defining the mask were identical to the motion trajectories of the walker points. Under these conditions, subjects generally performed at chance levels in the short trial duration condi-tion. In the long trial duration condition, per-formance was generally above chance and depended upon ISI. Above-chance perform-ance with ISIs greater than 50msec is thought to reflect high-level motion processes (Anstis, 1980; Baker & Braddick, 1985). Such processes may allow for attentional tracking of the point-light walker over extended temporal intervals (Cavanagh, 1992; Lu & Sperling, 1995; Thorn-ton, Rensink, & Shiffrar, 1998). Interestingly, neural representations of action are influenced by attentional processes (Decety, 1996). How-ever, other aspects of the results of this experi-ment cast serious doubt on the hypothesis that the visual perception of human movement depends exclusively on high-level neural processes. First, in the long trial duration con-ditions, performance was at ceiling when ran-dom dot masks were used but significantly below ceiling when scrambled walker marks were employed. Since scrambled walker masks effectively eliminate the utility of local motion analyses, suboptimal performance with these masks can be attributed to the loss of input from local analyses. Second, in the long trial duration condition of Experiment 2, subjects could not accurately judge the walker's direction at long ISIs. This finding further supports the importance of temporally restricted, or low-level motion analyses. Thus, the results of these experiments suggest that both local and global processes contribute to our visual interpretation of the movements of others.

Since low-level motion detectors may serve as the gateway to the perception of object motion, it might not be surprising that they

play an important role in the visual perception of human movement. Indeed, models involving strictly local computations do capture some aspects of the visual perception of human movement (Hoffman & Flinchbaugh, 1982; Webb & Aggarwal, 1982). However, such approaches cannot explain the orientation specificity (Ahlström et al., 1997; Bertenthal & Pinto, 1994; Pavlova, 1989; Sumi, 1984) nor the spatio-temporal limits within which we can visually identify a moving human. It is also unclear how such models can be extended to account for our ability to visually classify different human actions (Dittrich, 1993; MacArthur & Baron, 1983). Thus, the critical question becomes, what is the nature of the high level mechanism(s) involved in the visual perception of locomotion? Neurophysiological and case studies suggest that area STP may play an important role in the visual perception and/or interpretation of human movement (McLeod et al., 1996; Oram & Perrett, 1994; Perrett et al., 1990; Vaina et al., 1990). Since this region receives convergent input from the dorsal and ventral pathways (Baizer et al., 1991), it may be involved in the integration of form and motion cues (Perrett et al., 1990). This integration may contribute to the visual perception of a moving human form across space and time.

Another line of research suggests that the visual perception of human movement may involve a functional linkage between the perception and production of motor activity (Viviani, Baud-Bovy, & Redolfi, 1997; Viviani & Stucchi, 1992). In other words, the perception of human movement may be constrained by knowledge of human motor limitations (Shiffrar, 1994; Shiffrar & Freyd, 1990, 1993). Given our extensive visual exposure to people in action, it is possible that this implicit knowledge may be derived from visual experience. However, physiological evidence increasingly suggests that motor experience may be crucial to this visual process. For example, "mirror" neurons in monkey premotor cortex respond both when a monkey performs a particular action and when that monkey observes another monkey or a human performing that same action (Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). Recent imaging data clearly suggest that, in the human, the visual perception of human movement involves both visual and motor processes. That is, when subjects are asked to observe the actions of another human so that they can later imitate those actions, PET activity is found in those brain regions involved in motor planning (Decety et al., 1997). Thus, visual observation of another individual's movement can lead to activation within the motor system of the observer.

Interestingly, action observation without the intent to imitate does not consistently engage motor planning areas (Decety et al., 1997). Intentionality is known to play a fundamental role in the production of human movement (Bonnard & Pailhous, 1991, 1993; Laurent & Pailhous, 1986). Indeed, intentionality or the ability to actively modify muscle activity, marks the critical difference between animal and object movement. Since intentionality controls both the motor production and visual analysis of human movement, it may serve to connect the two processes. This proposed link-

age is consistent with the hypothesis that the perception of human movement may differ from the perception of other complex but non-intentional, motions. Taken together, these intriguing results suggest that we may understand the actions of others in terms of our own motor system. The high-level visual mechanism suggested by the results of the current behavioural experiments may well reflect this linkage between the visual and motor systems.

REFERENCES

Ahlström, V., Blake, R., & Ahlström, U. (1997). Perception of biological motion. *Perception, 26,* 1539–1548.

Anstis, S.M. (1980). The perception of apparent movement. *Philosophical Transactions of the Royal Society of London, 290,* 153–168.

Baizer, J., Ungerleider, L., & Desimone, R. (1991). Organisation of visual inputs to the inferior temporal and posterior parietal cortex in macaques. *Journal of Neuroscience, 11,* 168–190.

Baker, C., & Braddick, O. (1985). Temporal properties of the short-range process in apparent motion. *Perception, 14,* 181–192.

Bertenthal, B.I., & Pinto, J. (1994). Global processing of biological motions. *Psychological Science, 5,* 221–225.

Bonnard, M., & Pailhous, J. (1991). Intentional compensation for selective loading affecting human gait phases. *Journal of Motor Behaviour, 23,* 4–12.

Bonnard, M., & Pailhous, J. (1993). Intentionality in human gait control: Modifying the frequency-to-amplitude relationship. *Journal of Experimental Psychology: Human Perception and Performance, 19,* 429–443.

Burt, P., & Sperling, G. (1981). Time, distance, and feature trade-offs in visual apparent motion. *Psychological Review, 88,* 171–195.

Cavanagh, P. (1991). Short-range vs. long-range motion: Not a valid distinction. *Spatial Vision, 5,* 303–309.

Cavanagh, P. (1992). Attention-based motion perception, *Science, 257,* 1563–1565.

Cavanagh, P., & Mather, G. (1989). Motion: The long and the short of it. *Spatial Vision, 4,* 103–129.

Cutting, J.E. (1978). A program to generate synthetic walkers as dynamic point-light displays. *Behaviour Research Methods & Instrumentation, 10,* 91–94.

Cutting, J.E., Moore, C., & Morrison, R. (1988). Masking the motions of human gait. *Perception and Psychophysics, 44,* 339–347.

Decety, J. (1996). The neurophysiological basis of motor imagery. *Behavioural Brain Research, 77,* 45–52.

Decety, J., Grezes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F., & Fazio, F. (1997). Brain activity during observation of actions: Influence of action content and subject's strategy. *Brain, 120,* 1763–1777.

Dittrich, W.H. (1993). Action categories and the perception of biological motion. *Perception, 22,* 15–22.

Green, M. (1986). What determines correspondence strength in apparent motion? *Vision Research, 26,* 599–607.

Heptulla-Chatterjee, S., Freyd, J., & Shiffrar, M. (1996). Configurational processing in the perception of apparent biological motion. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 916–929.

Hoffman, D.D., & Flinchbaugh, B.E. (1982). The interpretation of biological motion. *Biological Cybernetics, 42,* 195–204.

Inman, V.T., Ralston, H., & Todd, F. (1981). *Human walking.* Baltimore, MD: Williams & Wilkins.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics, 14,* 201–211.

Johansson, G. (1975). Visual motion perception, *Scientific American, 232,* 76–88.

Johansson, G. (1976). Spatio-temporal differentiation and integration in visual motion perception. *Psychological Review, 38,* 379–393.

Johansson, G., von Hofsten, C., & Jansson, G. (1980). Event perception. *Annual Review of Psychology, 31*, 27–63.

Korte, A. (1915). Kinematoskopische Untersuchunge. *Zeitschrift fuer Psychologie, 72*, 194–296.

Kozlowski, L.T., & Cutting, J.E. (1977). Recognising the sex of a walker from a dynamic point-light display. *Perception and Psychophysics, 21*, 575–580.

Kozlowski, L.T., & Cutting, J.E. (1978). Recognising the sex of a walker from point-lights mounted on ankles: Some second thoughts. *Perception and Psychophysics, 23*, 459.

Laurent, M., & Pailhous, J. (1986). A note on modulation of gait in man: Effects of constraining stride length and frequency. *Human Movement Science, 5*, 333–343.

Lu, Z.L., & Sperling, G. (1995). Attention-generated apparent motion. *Nature, 377*, 237–239.

MacArthur, L.Z., & Baron, M.K. (1983). Toward an ecological theory of social perception. *Psychological Review, 90*, 215–238.

Marey, E.J. (1972). *Movement*. New York: Arno Press & New York Times (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.

McBeath, M.K., & Shepard, R.N. (1989). Apparent motion between shapes differing in location and orientation: A window technique for estimating path curvature. *Perception and Psychophysics, 46*, 333–337.

McLeod, P., Dittrich, W., Driver, J., Perrett, D., & Zihl, J. (1996). Preserved and impaired detection of structure from motion by a "motion blind" patient. *Visual Cognition, 3*, 363–391.

Oram, M., & Perrett, D. (1994). Responses of anterior superior temporal polysensory (STPa) neurons to "biological motion" stimuli. *Journal of Cognitive Neuroscience, 6*, 99–116.

Pantle, A.J., & Petersik, J.T. (1980). Effects of spatial parameters on the perceptual organisation of a bistable motion display. *Perception and Psychophysics, 27*, 307–312.

Pavlova, M. (1989). The role of inversion in perception of biological motion pattern. *Perception, 18*, 510.

Perrett, 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, UK: Cambridge University Press.

Petersik, J.T. (1989). The two-process distinction in apparent motion. *Psychological Bulletin, 106*, 107–127.

Petersik, J.T. (1991). Comments on Cavanagh and Mather (1989): Coming up short (and long). *Spatial Vision, 5*, 291–301.

Pinto, J., Thornton, I.M., & Shiffrar, M. (1998). *Orientation effects on the temporal integration of biological motion displays.* Manuscript in preparation.

Poizner, H., Bellugi, U., & Lutes-Driscoll, V. (1981). Perception of American Sign Language in dynamic point-light displays. *Journal of Experimental Psychology: Human Perception and Performance, 7*, 430–440.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research, 3*, 131–141.

Shepard, R.N. (1984). Ecological constraints on internal representation: Resonant kinematics of perceiving, imagining, thinking, and dreaming. *Psychological Review, 91*, 417–447.

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., Lichtey, L., & Heptulla-Chatterjee, S. (1997). The perception of biological motion across apertures. *Perception and Psychophysics, 59*, 51–59.

Sumi, S. (1984). Upside-down presentation of the Johansson moving light-spot pattern. *Perception, 13*, 283–286.

Thornton, I.M., Rensink, R.A., & Shiffrar, M. (1998). *Active versus passive processing of biological motion.* Poster presented at the annual meeting of the European Conference on Visual Perception, Oxford, UK.

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.

Van Essen, D.C., & DeYoe, E.A. (1995). Concurrent processing in primate visual cortex. In M. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 383–400). Cambridge, MA: MIT Press.

Viviani, P., Baud-Bovy, G., & Redolfi, M. (1997). Perceiving and tracking kinesthetic stimuli: Further evidence of motor-perceptual interactions. *Journal of Experimental Psychology: Human Perception and Performance, 23*, 1232–1252.

Viviani, P., & Stucchi, N. (1992). Biological movements look constant: Evidence of motor-perceptual interactions. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 603–623.

Wallach, H. (1935). Uber visuell wahrgenommene Bewegungsrichtung. *Psychologische Forschung, 20*, 325–380.

Webb, J.A., & Aggarwal, J.K. (1982). Structure from motion of rigid and jointed objects. *Artificial Intelligence, 19*, 107–130.

Wertheimer, M. (1912). Experimentelle stuidien uber das Sehen von Beuegung. *Zeitschrift fuer Psychologie, 61*, 161–265.

Zeki, S. (1993). *A vision of the brain*. Cambridge: Cambridge University Press.