Sensitivity to motion features in point light displays of biological motion

SANDHITSU R. DAS *, MACIEJ T. LAZAREWICZ,
ROBERT C. WILSON and LEIF H. FINKEL
Department of Bioengineering, University of Pennsylvania, Philadelphia, PA 19104-6321, USA

Received 1 June 2006; accepted 20 July 2007

Abstract—Psychophysical experiments are described that measure the sensitivity to motion features in point light displays of biological motion. Three motion features were investigated: the relative motion of the thighs, the relative motion of the thigh and leg, and the velocity profile of the leg. The perceptual threshold for discriminating a change in each motion feature was compared in upright and inverted point light displays. We find that subjects are more sensitive to two of the motion features in the upright display configuration (relative motion of thighs, relative motion of thigh and leg), but more sensitive to the third feature (velocity profile of the leg) in the inverted configuration. We propose that perceptual sensitivity to features used in biological motion perception should be greater in upright versus inverted displays. The results suggest that motion features differ in salience in biological motion perception.

Keywords: Gait recognition; point-light display; cyclogram.

INTRODUCTION

Certain classes of objects are recognizable by their idiosyncratic motions — flickering candles, fluttering flags — and it has been proposed that these complex but characteristic motions are detected by specialized, expert visual subsystems (Cavanagh et al., 2001; Ullman, 1984). The perception of biological motion is a paradigmatic case of recognizing a class of objects (humans, animals) based on their distinctive spatiotemporal patterns of motion, and visual cortical areas specialized for biological motion processing have been identified in superior temporal sulcus (STS) and other cortical regions (Beauchamp et al., 2003; Grèzes et al., 2001; Grossman and Blake, 2002; Grossman et al., 2000; Oram and Perrett, 1996). The processing capabilities of the biological motion recognition system are

*To whom correspondence should be addressed. E-mail: sudas@seas.upenn.edu
most clearly demonstrated by the ability to perceive a form-in-motion from point-light displays (PLDs) (Johansson, 1973). In PLDs, each joint of an actor (shoulders, elbows, knees, etc.) is tagged with a small dot, and as the actor moves, only these moving dots are visible. On viewing these displays, the action is immediately apparent (walking, running, throwing a baseball) and to some degree, even figural attributes can be inferred such as gender or identity of the actor (Barclay et al., 1978; Cutting and Kozlowski, 1977; Loula et al., 2005; Troje, 2002) or his/her mood (Heberlein et al., 2004). As Johansson and many subsequent investigators have shown, it is possible to perceive biological motion in a PLD even when the actor’s points are embedded in a field of independently moving ‘noise’ points. Any single frame of a PLD embedded in noise contains minimal form information, thus object identification must involve detecting characteristic features of the motion.

The visual system is capable of analyzing motion at multiple levels in a point light display — the motion of individual points, local groups of points, or global holistic patterns of optic flow. Some evidence suggests that analysis of intermediate-level motion features is particularly important for recognition (Casile and Giese, 2005). In particular, the relative motion of limb segments distinguishes different gaits, and is an idiosyncratic signature of biological motion.

The visual system may preferentially use some intermediate-level motion features in discriminating biological motion — for example, the relative motion of thigh versus leg may be more useful in recognition than the motion of the foot versus the wrist. If so, recognition should be more sensitive to perturbations in these features that are directly used in biological motion recognition. While distortion in any motion characteristic will be visible if it deviates sufficiently from the norm, greater effects on recognition will result from distortions in the features used in recognition.

Sensitivities to different features cannot be compared directly: a 10% change in thigh angle may not be perceptually equivalent to a 10% change in leg velocity. Therefore, in the psychophysical experiments reported here, sensitivities are compared for changes in the same feature in upright versus inverted point-light displays. It is well established that perception of biological motion in inverted displays is more difficult (Grossman and Blake, 2001; Pavlova and Sokolov, 2000; Sumi, 1984). Upright walkers can be perceived when embedded in higher densities of random dots (Pavlova and Sokolov, 2000). Inversion impairs recognition of texture-defined walkers (Ahlström et al., 1997), recognition of the class of action (Dittrich, 1993) and recognition of the gender of the walker (Barclay et al., 1978). A similar effect of image inversion is found in face recognition (Farah et al., 1998) and other specialized, expert visual recognition systems (Bédard and Belin, 2004; Troje, 2003).

We hypothesize that if a feature is used by the visual system in biological motion perception (BMP), it will show an ‘upright-preference’ effect — namely, observers should be more sensitive to perturbations in that feature in upright stimuli compared to the same perturbation in an inverted stimulus. Features not explicitly used in biological motion perception (say for example, limb length) may not show
Sensitivity to motion features

an ‘upright-preference’. In fact, perturbations of features not explicitly used in biological motion recognition may be more noticeable in inverted displays. This might occur if upright displays are processed by an ‘expert’ biological motion system, which focuses preferentially on certain features; whereas inverted displays are processed by more general recognition systems that process all features more equally. If, as we will demonstrate, it emerges that changes in some features are more easily detected in upright displays whereas changes in other features are more easily detected in inverted displays, it suggests that not all features are processed similarly, but rather, certain key features may be of primary importance for recognition.

Given the large number of possible intermediate-level motion features, our choice of which few to psychophysically investigate is motivated by the gait analysis literature. These studies point to the relative motion of limbs as the key features for detecting subtle gait abnormalities (Charteris, 1982). Other studies have reported a relative insensitivity to the absolute velocity of limb motion (Viviani and Stucchi, 1992), suggesting that this feature is not critical for recognition. We present three psychophysical experiments that show a greater sensitivity, in upright displays, to changes in the relative motion of the thighs and to relative motion of thighs and legs; however, a greater sensitivity in inverted displays to changes in the velocity profile of the leg.

METHODS

Motion capture data

Gait data was obtained using the ReActor motion capture system (http://www.ascension-tech.com). Infrared emitters were placed at 13 body positions (head, shoulders, elbows, wrists, hips, knees, ankles) on 6 human subjects (3 male and 3 female). 3D spatial positions of markers were acquired at 33 frames/s with a spatial resolution of 3 mm. Subjects walked or ran at controlled speeds on a treadmill and cycled on a stationary bicycle. The ReActor system provides \((x, y, z)\) coordinates of the 13 markers. Line lengths, line angles, angular velocities and other variables are calculated from the data using the angles as defined in Fig. 1a. In the 2D image plane, the 13 dots have 26 degrees of freedom; alternatively, the 12 connecting lines can each be described by an angle and a length, and one reference dot, e.g., the head, is specified by an \((x, y)\) position, similarly yielding 26 degrees of freedom. In typical human walk, viewed in 90° profile, projected line (bone) lengths vary (up to 10%) over the gait cycle due to movement in the \(z\) (depth) plane.

An example of thigh and knee angle data acquired over two consecutive gait cycles is shown in Fig. 1b.
Figure 1. (a) $q$ is the thigh angle, defined with respect to vertical. $q > 0$ when the thigh is in front of the body, and $q < 0$ when the thigh is behind the body. $k$ is the knee angle, i.e., the angle of the leg with respect to the thigh. $k = 0$ at full extension and increases with knee flexion (Photograph by Eadweard Muybridge (1872), courtesy of the Archives of the University of Pennsylvania). (b) Graph showing the left and right thigh angles and the left knee angle over time for two consecutive gait cycles. The two thighs move in roughly antisymmetric fashion. During the forward swing of the thigh (e.g., ascending portion of red curve with circles) the thigh angle goes from negative to positive, and the knee first flexes (increasing angle) and then extends (magenta curve with squares). This figure is published in colour on http://www.ingentaconnect.com/content/vsp/spv

Psychophysics methods

Subjects. Eleven unpaid subjects (three female and eight male), including one of the authors (SD), participated in the study. Except for the author, the participants were uninformed about the purpose of the study. Four subjects (SD, SC, RR and JM) carried out all three experiments, two subjects participated in just two experiments and five subjects participated in just one experiment. Data from seven subjects were collected for each experiment. All subjects had normal or corrected-to-normal vision.

Stimuli. Point-light stimuli were generated from the motion capture data. The point-light walker was made of 13 point-lights marking the head, and the left and right shoulders, elbows, wrists, hips, knees and ankles. Each marker was displayed as a circular white dot on a black background. Each dot subtended a visual angle of $0.12^\circ$. The entire point-light walker subtended a visual angle of $1.6^\circ$ in the horizontal direction and $3.6^\circ$ in the vertical direction. All 13 point-lights were visible at all times during presentation of a walker — in other words, no occlusion cues were present.
The walker stimulus was presented on a 21′′ Sony Trinitron CRT monitor with a refresh rate of 120 Hz by a Macintosh G4 PowerMac computer using the Psychophysics Toolbox (Brainard, 1997) in MATLAB® (The Mathworks, Natick, MA). Each trial consisted of three stimulus intervals with an inter-stimulus interval of 700 ms. Each interval contained a 330 ms video clip of a point-light walker. A fixation cross remained visible throughout the duration of a trial in the middle of the screen. The walker was presented at a random location on each interval, the position with respect to the fixation cross varying on each stimulus interval between 0 and 1.5° of visual angle in both horizontal and vertical directions. This region falls within central vision in which the BMP inversion effect is found (Ikeda et al., 2005).

Task. Observers performed a 3-interval 2-alternative forced choice (2AFC) task. On each trial, either one or two of the three walker video clips was distorted by altering a certain motion characteristic (identical distortions were used when two intervals were distorted). The middle (second) interval contained either an undistorted or a distorted walker (chosen randomly). One of the flanking intervals (first or third), again chosen randomly, contained the distorted walker; the other interval contained the undistorted walker. In addition, the walker in the middle interval was always reflected horizontally so that direction of walking was reversed compared to those in the 1st and 3rd intervals. Using the middle interval as the reference, the task was to identify which one of the two flanking intervals contained the walker whose motion was different from the reference walker. This means that for the trials in which the middle interval contained the undistorted walker, the correct answer would be to choose the interval that contained the distorted walker. For trials in which the middle interval contained the distorted walker, the correct answer would be to choose the interval that contained the undistorted walker. Observers were asked to indicate their choice by pressing one of two keys on the keyboard, then he or she could press the Space Bar to move on to the next trial. The task does not determine whether the subject recognized biological motion in the display; rather, it measures ability to detect small changes in particular features. The 3-interval design was chosen so as to avoid any requirement to judge which stimulus appeared to display a more ‘normal’ walk, as such judgments are ill-defined. Since the task requires determining which intervals contain stimuli that are similar — the 1st and 2nd intervals, or the 2nd and 3rd intervals — the direction of walking was reversed in the middle interval to make more difficult any attempts to use a strategy based on local point matching. The task also requires the subject attend to all three intervals.

Procedure. In order to determine the threshold level of distortion in the specified motion characteristic that can be detected by the visual system, a 2-up 1-down staircase procedure was used. The staircase started at a high level of distortion, easily detectable by all observers. After two consecutive correct answers the
distortion level was lowered (i.e., the difficulty increased) by a factor of 0.26 (0.1 on a logarithmic scale). After an incorrect answer, the distortion level was raised by the same amount. For example, the relative motion of the thighs was distorted by delaying the motion of one thigh by a fixed time. The maximum distortion used corresponded to 50% of the gait cycle. Each staircase change in distortion level altered the time delay by 26% of its current value. The staircase procedure was terminated after 13 reversals. During pilot versions of the experiments, we observed periods of long sequences of consecutive correct responses alternating with long sequences of consecutive incorrect responses. We suspect that these large swings in staircase values were due to shifts in attentional or perceptual strategy (Thornton et al., 2002). By manipulating fine points of the stimulus we were able to obtain an experimental protocol where such swings were minimal, yet performance still consistently diminished at the end of the sessions for many subjects. Therefore, we discarded the last two reversals for all sessions. The threshold level of distortion was calculated as the average of the last 8 reversals. Observers were given two series of staircases as practice sessions. In addition, before each staircase session, observers went through a few practice trials at the highest distortion level. During these trials, the clips were longer (750 ms) and trial-by-trial feedback was given. No feedback was given during the actual experiments. For each experiment, separate staircases were used for upright walker stimuli and inverted walker stimuli. The order in which observers carried out the different experiments and upright vs. inverted staircases was counterbalanced across observers.

RESULTS

Motion features in cyclograms

In clinical applications of gait analysis, the cyclogram has proven to be a representation that is very useful for analyzing gait abnormalities (Charteris, 1982; Goswami, 1998; Grieve, 1969). The cyclogram is simply a plot of one joint angle or joint velocity against another over time (see Figs 2 and 3). Each point on the curve represents the ‘body form’ at a single time point; the 2D shape of the curve represents the motion. The cyclogram is a parametric curve — it renders time implicit; it also makes implicit information on limb length, step size, and limb acceleration.

Figure 2 shows the angle of the left thigh plotted against the angle of the right thigh over 9240 frames of walking videos recorded from 6 subjects across speeds ranging from 3–5 mph. The data show that the two thigh angles are anti-correlated — the thighs swing in opposite directions in a roughly symmetrical manner. However, as the thigh reaches the end of its forward ‘swing’, it decelerates in preparation for transfer of body weight (note leveled-off crests of the thigh angle curves in Fig. 2).

Figure 3 plots thigh angle versus the ipsilateral knee angle for one gait cycle of a walking and a running video. The cyclogram has a characteristic shape — we will refer to this shape as a ‘D’. The ‘D’ is traversed once in the clockwise direction
Sensitivity to motion features

Figure 2. (a) Scatter diagram of 9240 frames of gait cycles showing left thigh angle vs. right thigh angle across subjects. The best fit line to the middle portion of the swing (−20–20°) is plotted in yellow (slope of −1.01). (b) Scatter diagram of the same frames as in (a) showing left vs. right thigh angle velocity. The best fit line to the middle portion of the swing is shown in yellow as in (a) (slope of −0.95). Color represents phase of gait cycle. This figure is published in colour on http://www.ingentaconnect.com/content/vsp/spv

Figure 3. “D” figure showing thigh angle vs. knee angle over one gait cycle for walk and run. The “D” is traversed once in the clockwise direction during each gait cycle. The semicircular arc of the “D” represents the swing phase and the straighter vertical segment corresponds to the pivot phase. This figure is published in colour on http://www.ingentaconnect.com/content/vsp/spv

on each gait cycle. The semicircular arc represents the forward swing — note the thigh angle increases from about −20° to ∼20°. During the first half of the swing (thigh angle <0, i.e., behind the body), the knee angle increases (e.g., flexion). Shortly after the thigh crosses the vertical (thigh angle = 0), the knee angle begins to decrease (extension). The vertical segment of the ‘D’ represents the pivot phase, as the thigh pivots from ahead of the body to behind. During this pivot phase, the knee angle remains relatively constant (extended to bear weight). The cusp at the lower left of the ‘D’ is the ‘double support’ phase in which weight is transferred and the power stroke begins. The left and right limbs traverse the ‘D’ with an 180° phase lag. As one limb begins the swing, the other begins the pivot. Each dot plotted in Fig. 3 represents data on one video frame. Note the lower density of dots during the swing versus the pivot — the limb moves faster during the swing. Since the angles
traversed during swing and pivot are similar \( (\pm 20^\circ) \), the swing phase takes slightly less overall time. This is compensated for by the double support phase.

Together, the cyclograms in Figs 2 and 3 capture the basic elements of walking. The features represented include (1) antisymmetric scissors-type motion of the thighs, (2) a ‘swing phase’ characterized by knee flexion during the first half of the swing, followed by knee extension during the second half of swing, and (3) a ‘pivot phase’ in which the knee remains relatively straight and extended. These features define the relative motions of the two thighs and the relative motions of the thigh versus the leg. Figure 4 shows the robustness of the motion feature describing the relative movement of the thigh and the knee. The top panel plots the time at which knee flexion begins with respect to the forward swing of the thigh. Over 90\% of the time, knee flexion starts in the first 10\% of the swing — the mean onset time for knee flexion (for walking videos) is at 10.25\% of the swing with a standard deviation (SD) of 4.5\% (for running, the values are 5.61 \( \pm \) 3.3\%). Thus, the start of knee flexion is very tightly bound to the start of swing. Similarly, the end of knee extension is very tightly correlated with the end of swing. The mean and SD of the swing phase at which knee extension terminated is 96.74 \( \pm \) 4.18\% (walk) and 84.32 \( \pm \) 6.82\% (run). The transition from flexion to extension always happens at

![Figure 4](http://www.ingentaconnect.com/content/vsp/spv)
Sensitivity to motion features

Figure 5. Robustness of the “D” figure. Left panel shows the variability of the “D” figure across different gait cycles for the same subject walking at the same speed. Middle panel shows the variability of the “D” figure across different speeds of walking for the same subject. Right panel shows the variability of the “D” figure across different subjects walking at the same speed. This figure is published in colour on http://www.ingentaconnect.com/content/vsp/spv

or slightly after the midpoint of the swing. This event has more variability than the start of flexion or the end of extension. For walking, it varies between 40% and 80%, with the mean and SD being 62.23 ± 9.77%. For run, the histogram is sharper, with a mean of 77.22 ± 5.71%. Borghese and colleagues (Borghese et al., 1996) have similarly shown invariance of the time course of movements over the gait cycle across subjects.

Figure 5 shows the variations in the thigh–knee cyclogram across subjects. The left panel shows the same subject, walking at the same speed, for three subsequent gait cycles. The middle panel shows slight differences that emerge as the subject walks at different speeds. The right panel compares three subjects all walking at the same speed. Despite slight variations, the basic characteristics of the cyclogram are preserved: the ‘D’ shape of the curve, and the approximate position of the curve with respect to thigh and knee angles. Variations are to be expected based on differences in body size, and the individual style of gait.

The utility of the cyclogram representation for clinical gait analysis suggests that it captures key features of human gait. This suggests that gait abnormalities, or distortions, that result in significant changes in the cyclogram may be more perceptible than distortions that leave the cyclogram relatively unchanged. In other words, those motion features explicitly represented by the cyclogram (mainly relative motion of limbs) may correspond to the features to which the visual system is most sensitive and which it preferentially employs for recognition. Features that are implicitly represented in the cyclogram, or not represented at all, may be of lesser salience for recognition.

A second motivation for choosing relative limb motions as key features comes from principal component analysis (PCA) of gait. Troje and colleagues have found that PCA yields gait representations that aid gender recognition (Troje, 2002). Casile and Giese (2005) have also shown that critical features for biological motion recognition can be extracted from PCA. We carried out PCA on a dataset of 6 gait variables: left and right thigh angle, left knee angle, left and right thigh velocities, and left knee velocity. PCA revealed three components that accounted for 91%
of the variance. These components corresponded to the relative angle between the thighs, and the thigh–knee angle during swing and pivot (Das et al., 2006). Although previous studies have found correspondences between PCA and related reduced-dimensionality representations and receptive field properties in the visual system (Bell and Sejnowski, 1997; Olshausen and Field, 1996; Tailor et al., 2000), emergence of relative limb motions as principal components in the data in no way guarantees that these are the features used by the visual system. However, it does signify that these relative motion features, which are explicitly represented by the cyclogram, describe the data in an efficient manner, and would be useful as features for recognition.

Based on these considerations, and our assumption that sensitivity to features used for biological motion recognition will be greater in upright displays, we hypothesize that gait distortions that result in significantly distorted cyclograms will be more easily detectable in upright displays, whereas gait distortions that leave the cyclogram relatively intact may be more detectable in inverted displays. Three psychophysical experiments were carried out to test this hypothesis.

Experiment 1 — sensitivity to thigh antisymmetry

In this experiment, the antiphase symmetrical motion of the thighs was perturbed by delaying the movement of one thigh with respect to the other (Fig. 6d). Motion of the lower leg with respect to the thigh was not altered. The maximum distortion used was a time delay of 300 ms, which corresponds to 50% of the average gait cycle at walking speeds of 3 miles/hour.

Results. This experiment tested observers’ sensitivity to distortions in the relative phase relation between the two thighs. Perceptual thresholds for discriminating distortions in the upright walker were lower than those for the inverted walker for all seven observers. For example, the threshold distortion level for the upright condition for observer SD was 51 ms delay of the swinging leg with respect to the pivoting leg. This corresponds to \(~5\%) of the total duration of the average gait cycle in our stimuli — or a \(~10\%) deviation from the normal antiphase symmetry in the relative motion of the two legs. In contrast, the perceptual threshold for the inverted walker was a 96 ms delay, or an \(~21\%) deviation from the normal antisymmetric relationship between the two legs. This means that observer SD was roughly twice as sensitive to distortions of this motion feature in the upright condition compared to the inverted condition. For different observers, there was considerable variability in absolute threshold levels (51–198 ms for upright, 96–267 ms for inverted), yet for all observers, the threshold for the inverted condition was at least 50% higher than that for the upright condition, and in each case, the difference in threshold was statistically significant \((p < 0.01\) for all observers). A paired \(t\)-test across all observers showed significantly lower threshold in the upright condition \((t(6) = −6.2110, p < 0.001)\). A non-parametric Wilcoxon signed rank test con-
Sensitivity to motion features

Figure 6. Effect of distortions used in psychophysics experiments as revealed in the cyclograms. (Top) Effect of distortions on thigh-vs.-knee angles as shown in “D-shaped” cyclogram. (Bottom) Effects of distortions on antiphase symmetry between thighs. All panels show plots of the undistorted parameters vs. distortion at the minimum threshold level for subject RR. The mean square displacement $d$ between undistorted and distorted cyclogram appears at the top right corner of each panel. Panels (a) and (d) show distortions introduced in the thigh anti-symmetry experiment. Distorted stimuli alter the anti-symmetry between the two thighs (d), while maintaining the “D-shape” relation between thigh and leg (a). Panels (b) and (e) show distortions introduced in the experiment in which the amount of knee flexion and extension was reduced. The “D-shaped” cyclogram is significantly distorted, while thigh anti-symmetry is undisturbed. Panels (c) and (f) show the distortions introduced in the experiment in which the velocity profile of the thigh and leg was altered. In this case, there are only minor variations to “D-shaped” cyclogram and to the anti-phase symmetry of the thighs. Thus, in both the anti-symmetry and flexion-extension experiments, one cyclogram undergoes a significant amount of distortion, while in the velocity experiment, both cyclograms undergo a smaller change—leaving the major characteristics of “D-shape” and anti-symmetry intact. This figure is published in colour on http://www.ingentaconnect.com/content/vsp/spv

Confirmed these results ($p < 0.02$). The mean difference in threshold between the upright and inverted conditions was $57 \pm 24$ ms.

Experiment 2 — sensitivity to thigh–leg relative motion

In this experiment, the flexion and extension of the knee was reduced in the swinging leg (Fig. 6b). The maximum level of distortion corresponded to maintaining a constant knee angle of $0^\circ$ such that the leg did not flex or extend at all over the gait cycle.
Results. This experiment tested observers’ sensitivity to the changes in flexion and extension of the walker’s knee during the swing phase of gait. All seven observers displayed greater sensitivity to distortions in the upright walker condition than in the inverted walker condition. The average range of knee flexion/extension across our stimuli was 50°. For observer SC, the threshold for detecting a distortion in upright displays was an 8.5° change in angle range. This corresponds to 17% of the normal range of motion. In inverted displays, the threshold distortion was 20.5°, corresponding to 41% of the normal range of motion. Thus, observer SC was more than twice as sensitive to changes in this motion feature in the upright condition than in the inverted condition. The absolute threshold levels varied considerably among all observers (8–35° for upright, 20.5–44.5° for inverted); however, threshold for the upright condition was significantly lower than threshold for the inverted condition ($p < 0.01$) for all observers. In fact, with the exception of one observer (MR), the threshold levels for the upright condition were at least 30% lower than that for the inverted condition for all observers. A paired $t$-test across all observers showed significantly lower threshold in the upright condition ($t(6) = -6.4751, p < 0.001$). A non-parametric Wilcoxon signed rank test confirmed these results ($p < 0.02$). The mean difference in threshold between the upright and inverted conditions was 11 ± 4.5°.

Experiment 3 — sensitivity to velocity profile

In this experiment, distortion was introduced in the velocity profiles of the legs. Segments of the gait cycle were identified in which, for both left and right lower limbs, both the thigh and leg sustained monotonic angular movement (clockwise or anticlockwise). Within these segments, the angular velocity profile of both lower limbs was multiplied by a Gaussian envelope (Fig. 6c and 6f) so that during these segments, the lower limb (thigh and leg) initially moved more slowly than normal, then faster than normal in the middle portion of the segment, and then again slower than normal towards the end of the segment. The Gaussian multiplication is normalized so that the total angular displacement of the limb was not changed by this procedure. Both limbs (left and right) were distorted. The maximum level of distortion corresponded to a peak velocity of 2.8 times the original.

Results. This experiment tested observers’ sensitivity to a change in the angular velocity profile of the legs while retaining, to a large degree, the symmetry of relative movement between the thighs and the characteristic flexion and extension of the knee. Five out of seven observers were more sensitive to these changes in the inverted condition than in the upright condition. For the remaining two observers (JM and RH), the sensitivity under the two conditions was not significantly different. As an example, for observer RR, the normalized threshold for the upright condition corresponded to an increase of 113% in the peak angular velocity of the lower leg. Threshold for the inverted condition was an increase of 33% in the peak angular velocity — significantly lower than the upright condition ($p < 0.01$). For three
other observers, the threshold for the upright condition was at least 50% higher than that for the inverted condition, and for one observer, the difference was about 25%. In general there was considerable variability in the absolute sensitivities across observers (an increase of 72–149% in the peak velocity of lower leg for upright, and 33–100% for inverted). A paired t-test across all observers showed significantly higher threshold in the upright condition ($t(6) = 3.8872, p < 0.005$). A non-parametric Wilcoxon signed rank test confirmed these results ($p < 0.02$). The mean difference in threshold between the upright and inverted conditions was $37 \pm 25\%$ increase in peak velocity.

Some of the observers had more practice with point-light displays than others, because they had participated in earlier pilot versions of the experiments. It is known that performance in biological motion perception can improve considerably with practice (Grossman et al., 2004). This may be one reason the threshold values varied across observers — yet, the inverted-preference effect, that is, the group difference in threshold between upright and inverted conditions was present in all experiments (significant $p$-values across observers in all experiments).

**Effect of the distortions on the cyclogram.** The distortions used in the three experiments are shown in Fig. 6. The top panels show the normal and distorted gaits in the thigh–knee cyclogram; the bottom panels show the normal and distorted gaits in the thigh–thigh cyclogram. One way to quantify the amount of distortion in each experiment is the mean square displacements (msd) between the undistorted and distorted cyclograms, as measured in each panel of Fig. 6. Note that the distortion used in Experiment 1 significantly alters the thigh–thigh cyclogram (msd = 7.88), but leaves the thigh–knee cyclogram relatively unchanged. Conversely, the distortion used in Experiment 2 significantly changes the thigh–knee cyclogram (msd = 24.19), but does not affect the thigh–thigh cyclogram. Experiment 3 caused a small change in both cyclograms, but the magnitude of these changes were small compared to the respective distortions induced in Experiments 1 and 2 (thigh–knee msd = 8.79 in Experiment 3 compared to 24.19 in Experiment 2; thigh–thigh msd = 1.23 in Experiment 3 versus 7.88 in Experiment 1). These findings are corroborated by direct measurements of the displacement distance between corresponding point lights (head, elbow, etc.) in the distorted versus undistorted displays, summed over a gait cycle. Both the total displacement over the gait cycle, as well as the maximum displacement on any individual frame, show that Experiments 1 and 2 involve much larger changes in point light position than Experiment 3. For example, in Experiments 1 and 2, the maximal displacement on any frame of a point light between distorted and undistorted displays is a 14 pixel displacement (Experiment 1) or a 16 pixel displacement (Experiment 2) of the left ankle. In comparison, the maximal displacement in experiment 3 is 5 pixels (also of the left ankle). Conversely, Experiment 3 involves a much larger change in the maximum velocity of a point light between the distorted and undistorted displays. The maximum difference in velocities of a point light is 19 pixels/s in Experiment 1,
and is 37 pixels/s in Experiment 2, compared to a difference of 82 pixels/s in Experiment 3.

Our results thus show that, for the three motion features investigated, changes that are manifested by significant distortions in body positions as reflected in the cyclogram representation are more sensitively detected in the upright configuration, whereas a change that results in only a small distortion of the cyclogram but larger changes in maximum point light velocity are more sensitively detected in inverted displays.

DISCUSSION

Object recognition depends upon detection of features in the stimulus that are salient with respect to categorization (Sigala and Logothetis, 2002). We have attempted to identify certain ‘motion features’, suggested by cyclogram representations of gait, that are salient for the perception of biological motion. Our strategy rests on the premise that subjects will be more sensitive to perturbations in features used for recognition than perturbations in features that are not used for recognition. Of course, even features that are not preferentially ‘used’ are still visible, but variation in these features does not affect the categorization.

Given our overwhelming everyday experience with upright humans, it is not surprising that many aspects of biological motion perception are diminished in inverted displays, and one might therefore expect less sensitivity to detecting perturbations of any feature in inverted displays. However, we found greater sensitivity to at least one feature (e.g., velocity profile of thighs) in inverted displays. The identification of a feature, the sensitivity to which is not diminished, but in fact accentuated by image inversion, is key to the argument that certain specific features are preferentially used in biological motion perception.

Our results support the idea that in upright PLDs, specialized visual circuits are activated, and are sensitive to small perturbations in their preferred stimuli. In inverted displays, these specialized circuits are not strongly activated and thus attention may be directed to more general perceptual processes (Hemeren, 2005) — these general processes may be more sensitive to discrimination of other features (such as velocity of individual points). Interestingly, prosopagnosic patients perform better at recognizing inverted faces than upright faces (Farah et al., 1995). And patients with periventricular brain damage tested on biological motion stimuli are impaired at recognizing upright walkers but not inverted walkers (Pavlova et al., 2005). This supports the idea that increased detection sensitivities in upright displays is a characteristic of the specialized, expert cortical recognition system, and that non-specialized perceptual mechanisms can show different sensitivities. Additional support for this idea comes from Grossman and colleagues (Grossman et al., 2005) who showed that repeated transcranial magnetic stimulation over the superior temporal sulcus leads to poorer performance in recognizing upright point-light walkers but does not affect recognition of inverted point-light walkers.
The psychophysical experiments are consistent with the hypothesis that rather modest changes in the stimulus are perceptually detected provided they affect the cyclograms (thresholds for upright displays in the first two experiments corresponded to a 17% change in the range of the parameter). In contrast, as shown in Experiment 3, relatively large perturbations to the stimuli can be undetected provided the changes do not affect the cyclogram (an 80% change in the parameter). A separate demonstration of this finding can be observed by replacing the normal, approximately sinusoidal angle changes in joint angles during gait (e.g., see Fig. 1) with a triangular wave approximation (piecewise linear between max and min points on the curve). Despite this drastic change, the resulting video appears as normal walking (Lazarewicz et al., 2004), and generates idealized-looking cyclograms. This observation supports the finding that the velocity profile is not salient for gait perception. Viviani and Stucchi (1992) showed that velocity judgments in biological motion can be significantly erroneous — subjects perceived velocities of point-lights as uniform even though the variation in velocity was in excess of 200%. These effects may stem from an underlying mis-estimation of object speed that arises as the orientation of the object deviates from the direction of motion (Georges et al., 2002; Seriès et al., 2002). Regardless, it suggests that, at least within the range of motions tested here, the velocity profile of the legs is not a useful feature for biological motion recognition. Interestingly, Cutting’s (1978) original algorithm for generating PLDs, produced gait motions that differ significantly from real gait data (see Fig. 7), yet the displays appear as normal gait. Cutting found that to generate realistic-looking gait, it was critical to have antisymmetrical motion of the thighs and compound pendular motion of the lower leg only during the swing phase. These features generate idealized-looking cyclograms, with a ‘D-shaped’ figure in the thigh–knee cyclogram and perfect antisymmetry in the thigh–thigh cyclogram. Thus, the cyclogram features we have investigated as potential motion features, were incorporated, by design, into these widely used routines.

An alternative explanation of our results might be that the subjects detect certain symmetries in the image — such as the motion of the two thighs — and these symmetries are preserved on inversion. Asymmetric features may be more significantly altered in appearance on inversion. Thus, inversion might predominantly highlight asymmetric features and the inversion effect might depend upon asymmetric features rather than features used in recognition. However, our results do not support this hypothesis, since the perturbations introduced in Experiments 1 and 3 were symmetrical, whereas the distortion in Experiment 2 was asymmetrical.

Instead, the results support the idea that the biological motion perception system is specialized for analyzing only certain types of perturbations to the basic body form. One prediction that follows is that whereas velocity distortions along the normal path of body motion in walking are not salient in the upright configuration (as velocities vary widely in different gait situations), a change in velocity of a point
Figure 7. Relationships between limb angles as calculated by Cutting’s algorithm (Cutting, 1978). (Top left) The relative motion of thigh and leg over the gait cycle follows an ideal “D” figure. The knee flexes during the first half of swing and then extends, while remaining perfectly straight during pivot. (Top right) The motion of the two thighs are perfectly antisymmetrical with the thigh angles always being equal and opposite of each other. (Bottom) The thigh angles and left knee angle over time capture the main features of Fig. 1(b), but differ significantly in detail (e.g., compare knee angle). This figure is published in colour on http://www.ingentaconnect.com/content/vsp/spv.

light in an unusual direction during gait would be at least as salient in the upright configuration as in the inverted configuration.

Necessity and sufficiency of motion features

The features considered here are most likely a small subset of those used by the visual system in biological motion perception. Bobbing of the head, motion of the pelvis, and even motions of the ankles contain robust information on gait. Several lines of evidence, however, provide support that the features investigated here are among those used by the visual system. These motion features are rarely found in non-biological motion. These features are diagnostic in differentiating types and styles of gait (see Fig. 5), and distinguishing other types of human motion. Most importantly, these features are consistent with the intuition that biological motion is defined by the relative motions of body parts. Early work (Bertenthal et al., 1984) showed that infants discriminate between upright and inverted point-light walkers based on the relative motion of groups of points. Studies of saccadic tracking of point-light displays show that observers focus on changes in orientations and
relative positions of the walker’s body parts as opposed to absolute motion of parts or translational motion of the body (Verfaillie et al., 1994). Booth and colleagues (Booth et al., 2002) showed that children are sensitive to the relative motion of two limbs as well as relative motion within a limb. Physiological studies (Jellema and Perrett, 2006) have shown cells in STS respond to knee flexion presented in isolation (without any other body movement).

The use of intermediate-level features is not a priori necessary for recognition, and Giese and Poggio (2003) have demonstrated successful recognition using a model based on global ‘snapshots’ of the instantaneous form and optic flow patterns of the display. In their model, neurons activated by successive snapshots are synaptically coupled via time-dependent synaptic plasticity, thus storing the sequence of snapshots. In the snapshot model, the global pattern of form and motion is captured at each time point, and the sequence of these ‘snapshots’ is then learned. There is no explicit characterization of the motion of individual limbs or body parts, and no difference in importance or weight is given to any feature of the stimulus — in fact, body features are not explicitly detected. The snapshot model is a general spatiotemporal recognition algorithm that could, given appropriate inputs, learn to recognize any spatiotemporal pattern.

Our results suggest that the visual system makes use of explicit information on relative limb motion. The shape, size, and position of the ‘D-shaped’ thigh–leg cyclogram, for example, convey detailed information identifying the motion and distinguishing it from other classes of motion. This information is not made explicit in the snapshot model, however, intermediate level features can be extracted from the snapshots (Casile and Giese, 2005). In addition, the snapshot model incorporates global form information which we have neglected. Evidence suggests both types of information converge on STS (Giese and Poggio, 2003).

The ability to recognize biological motion from form cues was previously demonstrated by Song and colleagues (Song et al., 2001). They showed that recognition of PLDs can be achieved by analyzing the statistics of configurations of triples of identified and ‘labeled’ body points over a video sequence. Since the spatial relations between identified body points change in a consistent and characteristic manner over the gait cycle, the presence of a set of these characteristic ‘triangles’ in a video sequence can be used to estimate the relative probability of biological motion in an image sequence. Our features can likewise be reduced to triples of points, but we focus on the motion (temporal sequence) of these triples. Song and colleagues’ results argue that knowing the sets of triples (i.e., the forms), without knowing the temporal ordering in which they occur (i.e., the motion) is sufficient to statistically identify the gait.

Additional support for extracting form at each time point comes from experiments by Lappe and colleagues (Beintema and Lappe, 2002; Lange and Lappe, 2006). They demonstrated biological motion perception in PLDs in which the position of the point light is moved randomly along the limb from frame to frame during the display. Although such displays retain significant motion information (Casile
and Giese, 2005), they make the extraction of structure-from-motion much more difficult. Lappe and colleagues’ findings are not inconsistent with our model: sequences of forms may be extracted at an intermediate rather than global level — for example in the processes occurring between V1 and areas MT and MST (Georges et al., 2002; Seriès et al., 2002).

More recently, Casile and Giese (2005) proposed that certain intermediate-level features may be critical for biological motion perception, and have described ‘opponent motion’ as one such feature. They constructed a PLD in which all dots move in random tracks except for those in two narrow horizontal panels, which move antisymmetrically in “x” — one panel corresponds to where the arms would be, the other where the legs would be. Observation of this display conveys an impression of a walker. This elegant demonstration shows that activation of a motion feature, in two appropriate spatial regions evokes BMP. It is important to note that point-light displays of walking subjects viewed frontally (head-on) can be seen as walkers, despite having almost no opponent motion between the dots. In the frontal view, it is likewise difficult to extract thigh–leg angles. However, an intermediate-level detector specialized for angle inputs might learn to integrate additional cues, such as the height of the ankles above ground. As shown in Fig. 8, frontally-viewed ankle motion is not strictly antisymmetric, but rather shows an alternating rise and fall, that could be learned/associated with other motion cues. In this manner, multiple other cues could be integrated, from the sounds of footfalls, to the ‘swish’ of nylon ski pants.

A number of authors have discussed possible parallels between the neural mechanisms for generating biological motion and mechanisms for perceiving it (Decety and Grèzes, 1999; Rizzolatti et al., 2001; Saygin et al., 2004). Mirror neurons in frontal cortex (Rizzolatti and Craighero, 2004) have been found to respond both to self-generated motor output and to the perception or mental envisioning of such motion. Perhaps most relevant to biological motion perception are findings on the movement abnormalities observed in limb apraxia — the loss of coordinated movement across limbs. Poizner et al. (1995) argue that apraxia is

![Figure 8. The y-coordinates (mm) of ankles in a frontal view of point-light walker for two gait cycles. In this case, movement in x is minimal, and movement in y has very little opponent motion. This figure is published in colour on http://www.ingentaconnect.com/content/vsp/spv](image)
characterized by (1) loss of proper joint synchronization, (2) inappropriate relative joint angles and (3) inappropriate phase angles between limbs. These motor abnormalities directly parallel the visual features most critical for biological motion perception, as captured in the cyclogram representations, and this commonality suggests that motor and sensory systems may represent movement in a similar fashion.

Acknowledgements

The authors are grateful to Dr. Benjamin Backus for guidance in the design of the psychophysical experiments and useful discussions. We thank the Center for Human Modeling and Simulation for use of the ReActor system, and David Kreiger for pointing out the apraxia literature.

This research was supported by the DoD Multidisciplinary University Research Initiative (MURI) program administered by the Office of Naval Research under grant N00014-01-1-0625. Commercial relationships: none.

REFERENCES


