Temporal "Bubbles" reveal key features for point-light biological motion perception

Steven M. Thurman

Department of Cognitive Science,
University of California, Irvine, USA

Emily D. Grossman

Department of Cognitive Science,
University of California, Irvine, USA

Humans are remarkably good at recognizing biological motion, even when depicted as point-light animations. There is currently some debate as to the relative importance of form and motion cues in the perception of biological motion from the simple dot displays. To investigate this issue, we adapted the "Bubbles" technique, most commonly used in face and object perception, to isolate the critical features for point-light biological motion perception. We find that observer sensitivity waxes and wanes during the course of an action, with peak discrimination performance most strongly correlated with moments of local opponent motion of the extremities. When dynamic cues are removed, instances that are most perceptually salient become the least salient, evidence that the strategies employed during point-light biological motion perception are not effective for recognizing human actions from static patterns. We conclude that local motion features, not global form templates, are most critical for perceiving point-light biological motion. These experiments also present a useful technique for identifying key features of dynamic events.

Introduction

Humans are extremely adept at recognizing the actions of others, even when the human form is represented by only a handful of tokens or "point-lights." Point-light biological motion animations depict complex human actions through joint kinematics, without explicit representation of body shape (Johansson, 1973). From these animations, observers can recognize a variety of actions such as walking, running, and dancing can identify the gender of an actor (Cutting, 1978) and recognize emotional expressions conveyed by body movements (Dittrich, Troscianko, Lea, & Morgan, 1996; Pollick, Paterson, Bruderlin, & Sanford, 2001).

Current computational models of biological motion perception differ in the extent to which they emphasize the importance of dynamic and structural information. That motion cues would be critical for point-light biological motion perception is intuitive. Take, for example, initial reports that describe single, static frames as a meaningless cloud of dots that emerge into a human figure only when set into motion (Johansson, 1973). Also disrupting the temporal ordering of point-light frames, or skipping frames altogether, impairs individuals' ability to recognize biological motion (Cutting, 1981; Mather, Radford, & West, 1992).

Point-light biological motion perception, however, is not disrupted by a number of manipulations that render local motion cues difficult to analyze. Point-light animations constructed with limited lifetime dots and dots that jitter spatial position from frame to frame should have minimal local image motion. And yet these displays are easily recognized as human actions (Beintema & Lappe, 2002; Neri, Morrone, & Burr, 1998).

What then is the relative importance of motion and form cues for point-light biological motion perception? To address this issue, we set about identifying the key features for recognition of biological motion using an adaptation of the "Bubbles" technique. Bubbles is a human observer-based feature extraction algorithm that identifies the diagnostic information used for image categorization tasks (Gosselin & Schyns, 2001). In this paradigm, randomly placed Gaussian windows partially reveal an image behind a dense mask. Observers make forced-choice discriminations and performance varies depending on the quality of the information revealed. Discriminations are most accurate when the sampled regions contain task-relevant information and less accurate when the critical information is inadequately sampled. Thus, based on observer performance, one can estimate the task-relevant, or diagnostic, regions of the stimulus space. The Bubbles technique is most commonly used to study the critical information in face discrimination (Gosselin & Schyns, 2001), object recognition (Gibson, Lazareva, Gosselin, Schyns, & Wasserman, 2007; Schyns & Gosselin, 2002), natural scene perception (McCotter, Gosselin, Sowden, & Schyns, 2005; Nielsen, Logothetis, & Rainer, 2006), and the perception of ambiguous images (Bonnar, Gosselin, & Schyns, 2002).

To identify diagnostic features in point-light animations, we have created temporal "Bubbles." In this paradigm, randomly selected intervals of biological motion are revealed within a motion–noise sequence. From the distribution of diagnostic information across the entire action sequence, we can deduce (1) whether some moments in biological motion are more diagnostic than others, and if so, (2) the features of point-light animations that correlate with these diagnostic moments. The results from these experiments dissociate three candidate key features for point-light biological motion perception, namely, global form, joint velocity, and opponent motion.

Temporal Bubbles

Methods

Five unpaid participants were recruited from the Psychology Research Participation Pool at the University of California Irvine. The participants had no prior experience viewing point-light animations and were naive to the purpose of the experiment. Participants gave written consent in accordance with the University's IRB protocol.

Stimuli were presented in Matlab using the Psychophysics Toolbox (Brainard, 1997; Pelli, 1997) on an Apple PowerPC G4 running Mac OS 9.2. Animations were viewed on a CRT (refresh 120 Hz) at a distance of 48 cm, which was maintained using a stationary chin rest.

The point-light walker was constructed with 13 dots representing the major joints and head of an actor, with all horizontal translational components eliminated making the walker appear to be on a treadmill. A full gait cycle consisted of 124 frames viewed at 60 Hz (inter-frame interval of 16.7 ms; approximately 0.5 Hz per cycle). Scrambled walkers were constructed by randomizing the starting position of the dots while leaving the motion vectors intact. The walker subtended 14×7 deg visual angle at the greatest extent in each direction. Point-lights were displayed as black dots (0.3 deg) against a white background. All animations were viewed within a rectangular stimulus window subtending 23×32 deg.

The temporal "Bubbles" stimuli were constructed as 3 s sequences of noise dots, into which a short target sequence (667 ms) was embedded (see Figure 1 and Movie 1). The motion trajectories of the noise dots were drawn from the biological sequence but with the spatial position randomized to perturb the hierarchical relations among the dots. At some point in the trial, a subset of the noise dots smoothly morphed into the target biological (or scrambled) walker, then morphed back into noise. To maintain "biological" trajectories during morphing, the noise dots that morph into target dots have trajectories computed backward from the desired target position. In other words, if a noise dot is destined to become the target ankle dot on frame 80, the inverse biological trajectory is applied to that ankle position to determine the *x, y* coordinate for frame 79 and so on. When played forward, the animation gives the appearance of a suddenly emergent biological figure, without disruptions in the smooth trajectories of the dots.

Because we are interested in how recognition waxes and wanes throughout the gait cycle, two temporal parameters
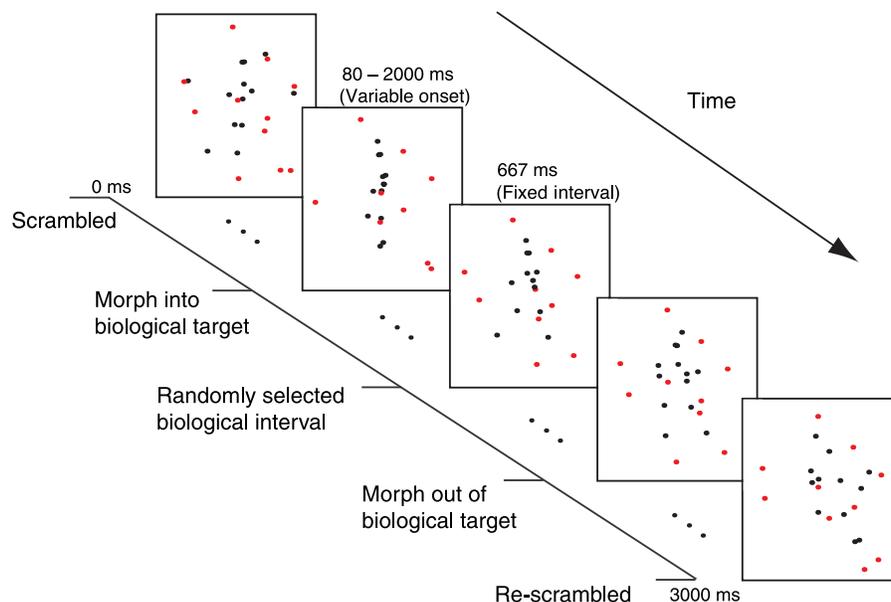

Figure 1. Schematic of a single trial. A subset of noise dots morph into a target animation (example shows a biological target) between 80 and 2000 ms after stimulus onset. Targets are displayed for 667 ms, after which those dots morph back into noise. For clarity, red dots represent noise dots, and black dots represent target (signal) dots. However, in the actual experiment all dots were black.
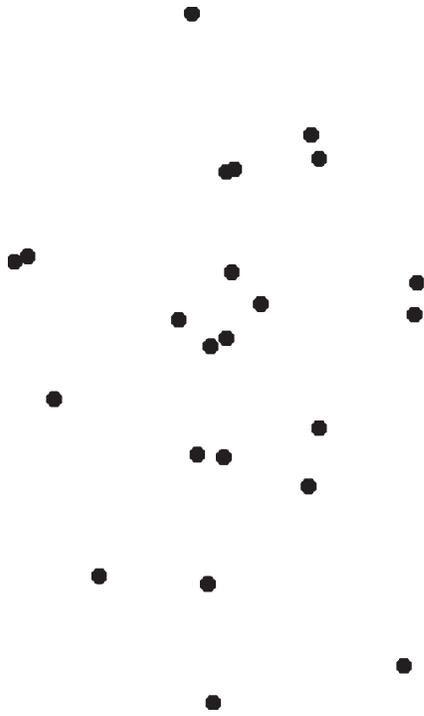
Movie 1. Sample of a single trial from the temporal "Bubbles" experiment. This 3-s movie includes 10 noise dots masking the signal dots. A leftward facing walker appears briefly in the middle of the sequence.

were randomized. First, the time at which the noise morphs into the target stimulus was randomized within an 80- to 2000-ms window after stimulus onset. By randomizing this stimulus onset window, we are assured that our recognition measures reflect stimulus specific effects and not more general fluctuations in sensitivity linked to the onset of the trial (e.g., some potential critical window after stimulus onset, fatigue, or wandering attention at the end of the 3-s trial). Second, the selected interval of the gait cycle inserted into the noise was randomly selected from trial to trial. This "temporal Bubble" is the measure of interest, and by randomizing the starting frame, we assure that each moment in the gait cycle has equal probability of being the first, middle, or last frame of the morphed target.

The experiment proceeded as follows. Participants were instructed to fixate on a central cross at the start of each trial. Participants viewed a single trial and made a forced-choice biological versus scrambled motion discrimination. To insure that single dots would not be predictive of stimulus type, target animations were randomly jittered ±4 deg around central fixation. To eliminate the possibility that subjects were making discriminations based on some figure-ground segregation (which might be less apparent in the scrambled walker), two additional participants repeated the Bubbles experiment discriminating the facing direction (leftward versus rightward) of the point-light walker.

Because we were interested in relative accuracy across the randomly selected moments in the gait cycle, we maintained overall accuracy, on average, at threshold (79%) using a double-interleaved 3-1 staircase. The number of noise dots masking the target animation was increased after three correct discriminations and decreased following one incorrect trial. Critically, the temporal moments revealed in the "Bubble" were randomly selected trial by trial and were not selected based on discrimination performance.

Prior to initiating the staircase, participants were shown five full cycles of the target walking sequence (which all observers spontaneously recognized as a human walker) and completed a practice block of the forced-choice task with the walker embedded in the dynamic noise. For the full experiment, subjects repeated the double-interleaved staircases three times, with the staircases terminating after about 20 reversals in accuracy (mean = 1028 trials, SD = 265). Those trials with a biological target were sorted into hits and misses, and those frames visible during that trial were scored accordingly. The temporal Bubbles solution is computed for each frame as the mean proportion correct (hits/hits + misses) when that frame is visible. This score measures how accurate the observers are when a given frame is shown in the target sequence.

Results

On average, observers required approximately 70 noise dots (SD = 20) to maintain discrimination performance at threshold level. This is similar to the noise tolerance thresholds for point-light biological motion reported in other studies (Grossman, Blake, & Kim, 2004; Hiris, Humphrey, & Stout, 2005).

All subjects revealed the same Bubble solution for the point-light walker whether discriminating biological from scrambled or leftward from rightward motion (Figure 2). The diagnostic moments in these biological animations wax and wane with an approximate sinusoidal pattern (mean phase shift = 6.6 frames, SD = 1.2). Evidently, biological motion sensitivity is not uniform across the gait cycles and is not task dependent. Both detection and discrimination are mediated by the same peak instances in the gait cycle.

Duration thresholds

To verify that subjects are indeed more sensitive to certain moments than others and to determine the minimum temporal window required for these critical moments, we measured duration thresholds for two key instances: diagnostic (peaks) and non-diagnostic (troughs). We argue that if the Bubble solution accurately reflects the use of critical information in these animations, then participants should be more accurate and require
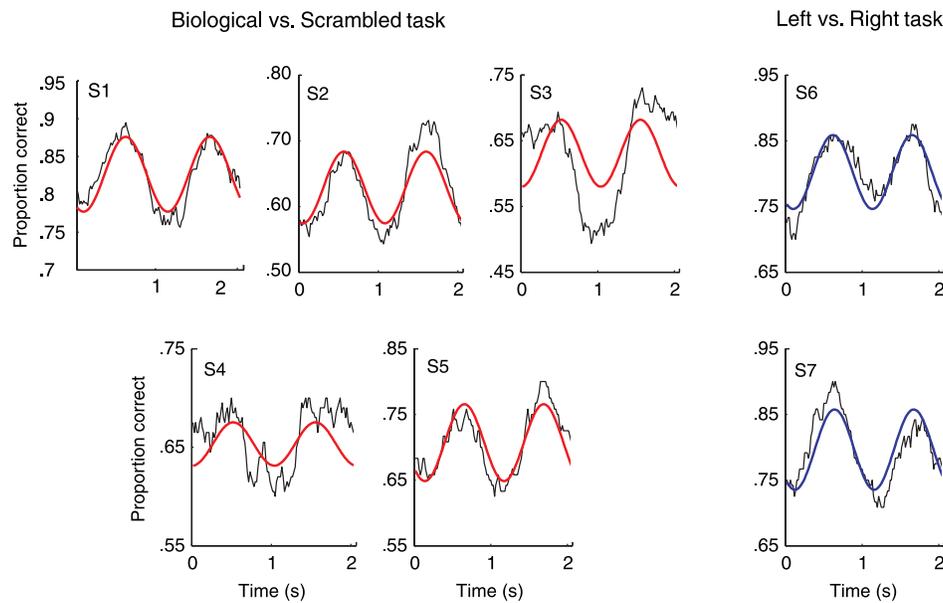
Figure 2. Individual subject performance for the biological versus scrambled task (left, $n = 5$) and left versus right task (right, $n = 2$). Performance is plotted as a function of time (124 total frames viewed over 2 s, although each trial only displayed a contiguous 667 ms). Observer accuracy for each frame (the Bubbles solution) is shown in black with the least-squares sinusoidal fit in blue and red.

shorter viewing intervals for discriminating instances of the diagnostic compared to the non-diagnostic intervals.

Methods

Four unpaid participants were recruited for this experiment. Experimental setup was the same as in Experiment 1. In this experiment, however, the target sequence was chosen either to be the most diagnostic (e.g., centered about Frame 38) or least diagnostic interval (centered about Frame 69, see Movies 2 and 3) based on the results from Experiment 1. Using the method of constant stimuli, we measured discrimination accuracy of the target sequences with 100, 200, 300, 400, or 500 ms duration (all shorter than the 667-ms exposure used in Experiment 1). Observers performed a forced-choice leftward versus rightward facing direction discrimination on each trial. To prevent ceiling or floor discrimination accuracy, animations were embedded in a fixed number of noise dots, determined individually from staircase estimates completed before the experimental blocks (mean noise dots = 40, $SD = 15$). Observers completed 50 trials per duration threshold (total of 500 trials), and mean proportion correct for each duration condition was computed. Threshold accuracy was estimated as the duration required for 75% discrimination performance based on the Weibull fit.

Results

Results from this experiment can be viewed in Figure 3. A within-subjects repeated-measures ANOVA revealed a

significant main effect of condition ($p < .045$) and time ($p < .046$; Greenhouse–Geisser correction for both). Participants were able to discriminate the facing direction of the walkers based on the diagnostic frames with an average of 294 ms ($SD = 78$ ms) of viewing. In contrast,
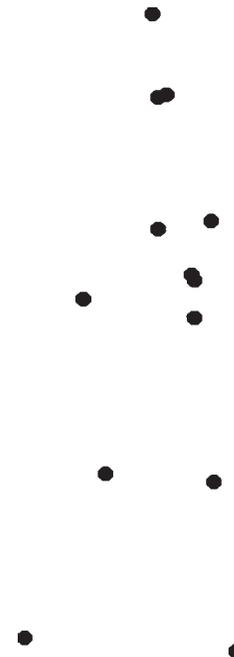
Movie 2. Movie of diagnostic frames of a point-light walker. Movie shows 500 ms centered on the peak (frame 38) of the Bubbles solution.
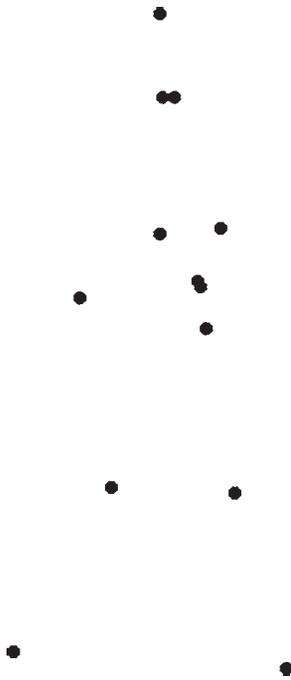
Movie 3. Movie of non-diagnostic frames. Movie shows 500 ms centered on the trough (frame 69) of the Bubbles solution.

four of five participants failed to reach threshold accuracy in discriminating the facing direction of the non-diagnostic animations, even with 500 ms of stimulus duration.

These results demonstrate the effectiveness of temporal Bubbles in identifying critical moments in a point-light action sequence. Observers are more sensitive (i.e. require shorter exposure) to the diagnostic moments in a point-light walker than to the non-diagnostic moments. We argue that these moments in the animation sequence are more perceptually salient because they contain a greater proportion of the key features useful for recognizing point-light biological motion. What are the key features that produce these diagnostic moments?

Analysis of diagnostic features

Experiments 1 and 2 revealed that sensitivity to point-light walker animations is highest at two moments in the gait cycle. What is happening at these times? Perhaps counter-intuitively, the Bubbles solution revealed the most diagnostic moments to be those instances in which the legs and the arms cross the midline, and the body as a whole is most vertically aligned (Movie 2). The least diagnostic moments are those with the most horizontal spread, that is, when the arms and legs are at their most extreme positions (Movie 3). Each of these moments occurs twice in a single gait cycle (i.e. left arm in front, then right arm in front).

There are at least two dynamic features of the point-light walker that correlate well with these diagnostic moments (Figure 4). First, performance is best when the extremities (e.g., wrists and ankles) move with the greatest joint velocity. Second, these high velocity moments also correspond to those instances in which the extremities cross, creating a local instance of opponent motion. In other words, performance is best when the distance between the extremities is least and they are moving in opposing directions. Thus, there is a significant correlation between the Bubble solution and the velocity and opponent motion profiles of the ankles (Pearson's $r = 0.67$ and -0.82, respectively; both $p < .001$). Note that opponent motion is the inverse of relative position, and thus when the ankles are far apart, local opponent motion will be weaker. We should also note that the velocity and the relative motion of the wrists are nearly identical to that of the ankles.

That these features in biological motion are correlated to discrimination performance is not surprising. Previous studies have identified body posture, local image velocity, and opponent motion as possible key features for point-light biological motion perception (Casile & Giese, 2005; Lange, Georg, & Lappe, 2006; Mather et al., 1992). Psychophysical studies have also noted the importance of the movement of the extremities (Mather et al., 1992), and in particular the feet (Troje & Westhoff, 2006), because they contain "ballistic-velocity" movement profiles that are unique to biological creatures.

The results from our Bubbles experiment would seem to imply, however, that the most diagnostic moments are those with the least apparent body structure. To directly
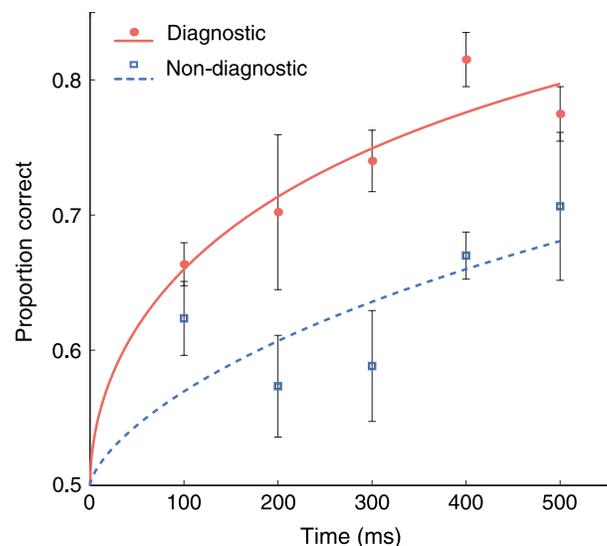
Figure 3. Mean duration performance ($n = 4$). Figure plots proportion correct as a function of the duration of the target stimulus in two conditions: diagnostic (red) and non-diagnostic (blue). Lines represent Weibull fits to mean data and bars ± one standard error of the mean.
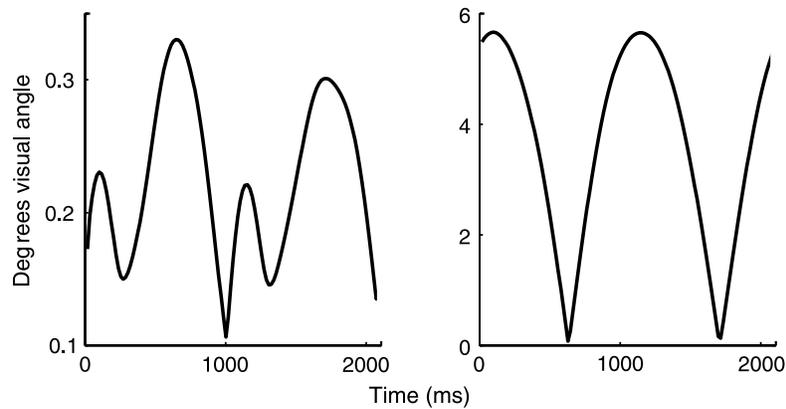
Figure 4. Left: instantaneous velocity over time. Figure plots the total Euclidian displacement of both ankles per frame. Right: relative motion of ankles over time. Figure plots horizontal distance between the ankle dots per frame (zero indicates legs crossing). Both figures are in terms of degrees visual angle.

test whether structural form cues or dynamic cues are primarily responsible for driving perceptual sensitivity to point-light biological motion, we measure sensitivity to single, static frames of the most and least diagnostic moments. If global form cues are driving the improved performance on the diagnostic frames, discrimination accuracy from an individual diagnostic frame should also be better than for a single frame from the least diagnostic moments. However, if the key features in biological motion perception are inherently dynamic, then observers should be equally good, or poor as it may be, at discriminating individual frames from the animation.

Methods

Five unpaid participants were recruited for this experiment. Experimental setup was the same as in previous experiments. Observers viewed 1.25 s trials of a single frame from the point-light walker sequence, embedded in a stationary noise array. Frames were either the most diagnostic frame (e.g., Frame 38), the least diagnostic frame, (Frame 69), or the frame corresponding to mean performance (Intermediate, Frame 53). Targets were either biological or scrambled and were jittered from fixation to prevent single dots from being indicative of trial type. Observers rated the trials as depicting either biological or scrambled, first in a training staircase with feedback until a stable threshold of noise dots was attained, then in three experimental blocks (totaling 600 trials, 100 samples of each condition) with the level of noise dots adjusted on-line with the same staircase procedure.

Results

Results from this experiment can be viewed in Figure 5. On average, observers required about 40 noise dots

(SD = 12) for threshold discrimination. A one-way ANOVA revealed a significant effect of condition ($p < .035$). To our surprise, observers were the least accurate discriminating a static instance of the so-called "diagnostic" interval (mean = 0.59, SD = 0.12). This frame was chosen because it corresponded to the moment in the walking animation when observers were most accurate. However, outside the context of the dynamic sequence, observers are least sensitive to this key instance. In fact, with stationary frames observers are most sensitive to the so-called "non-diagnostic" (mean = 0.74, SD = 0.06) and intermediate instances (mean = 0.75, SD = 0.07), which are structurally nearly identical.

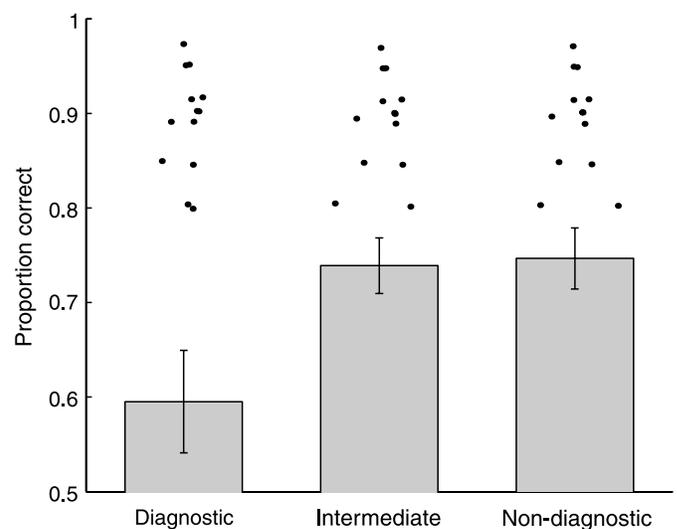Based on these findings, we conclude that posture is unlikely to be a key feature for the recognition of

Figure 5. Results of the static frame experiment. Bars represent mean accuracy (n = 5) in each condition. Error bars indicate ± one standard error of the mean. Point-light figures are corresponding sample stimuli for each condition.

biological motion in point-light animations. When viewed as stationary images, the pattern of sensitivity to the point-light walker sequence reverses such that moments of best performance are least diagnostic and vice versa. These results are strong evidence in favor of dynamic key features driving point-light biological motion perception.

Velocity versus opponent motion

Which dynamic feature is most critical for point-light biological motion perception? Both velocity and opponent motion of the extremities correlate well with perceptual sensitivity to a point-light walker. To dissociate these two features, we selected a new action, a jumping jack, in which joint velocity and opponent motion have unique signatures, particularly in the ankles (Figure 6). In jumping jacks, the ankles reach peak velocity twice per cycle (legs out, then in). Local opponent motion is highest (and relative motion lowest), however, only once per cycle (when the legs are close together).

Methods

Five unpaid, naive participants were recruited for this experiment. A point-light animation of two complete jumping jack cycles (118 frames, 32 ms inter-frame interval, 2 s per jumping jack) was selected for this experiment. The animation subtended approximately 18×12 deg, slightly larger than the point-light walker in Experiment 1. The experimental setup was otherwise the same as Experiment 1. Observers viewed 3 s Bubble animations, into which a 667-ms animation of a jumping jack (or scrambled jumping jack) was inserted. As in Experiment 1, noise dots were matched to the target animation by creating their movement vectors from the

vectors of randomly selected jumping jack dots. The number of noise dots was adjusted on-line with the same staircase procedure. Observers completed approximately 714 trials ($SD = 29$), discriminating the point-light jumping jack from scrambled jumping jacks.

Results

On average, observers required about 105 noise dots ($SD = 22$) to maintain threshold discrimination performance. All subjects revealed the same Bubbles solution for the point-light jumping jack (Figure 7). Due to the cyclical nature of the action, accuracy again varies with an approximate sinusoidal pattern (mean phase shift = -1.9 frames, $SD = 0.7$) that is significantly correlated to the relative distance between the ankles ($r = -0.57$, $p < .001$), but not the velocity ($r = 0.081$, $p = .38$). Based on these results, we argue that local opponent motion of the extremities is the primary critical feature for discriminating point-light biological motion.

Spatial Bubbles

The results from jumping jack Bubbles reveal opponent motion, but not local dot velocity or apparent body structure, to be the key feature driving sensitivity to point-light biological motion. As further evidence for this, we measured discrimination performance for point-light animations with and without local instances of opponent motion. In this "spatial Bubbles" experiment, we eliminated a randomly selected single dot from the point-light walker on each trial. Because local opponent motion is computed from the relative motion of two dots (particularly on the extremities: elbows, wrists, knees, and ankles), eliminating one of those dots should weaken that visual cue. Note that
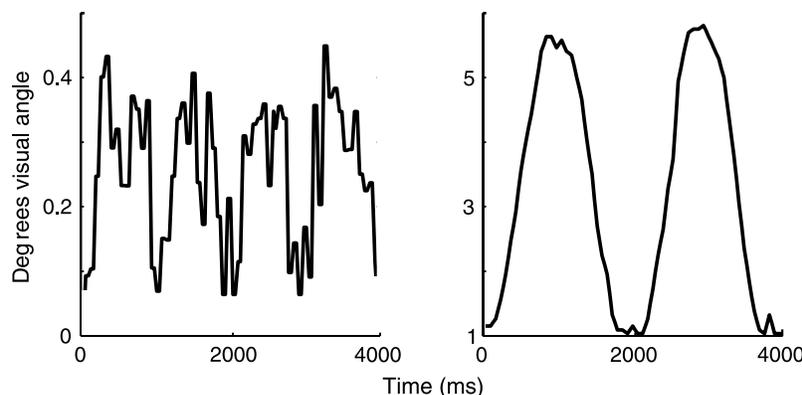

Figure 6. Left: Instantaneous velocity over time. Figure plots the total Euclidian displacement of both ankles per frame. Right: relative motion of ankles over time. Figure plots horizontal distance between the ankle dots per frame. Distances are plotted in degrees of visual angle.
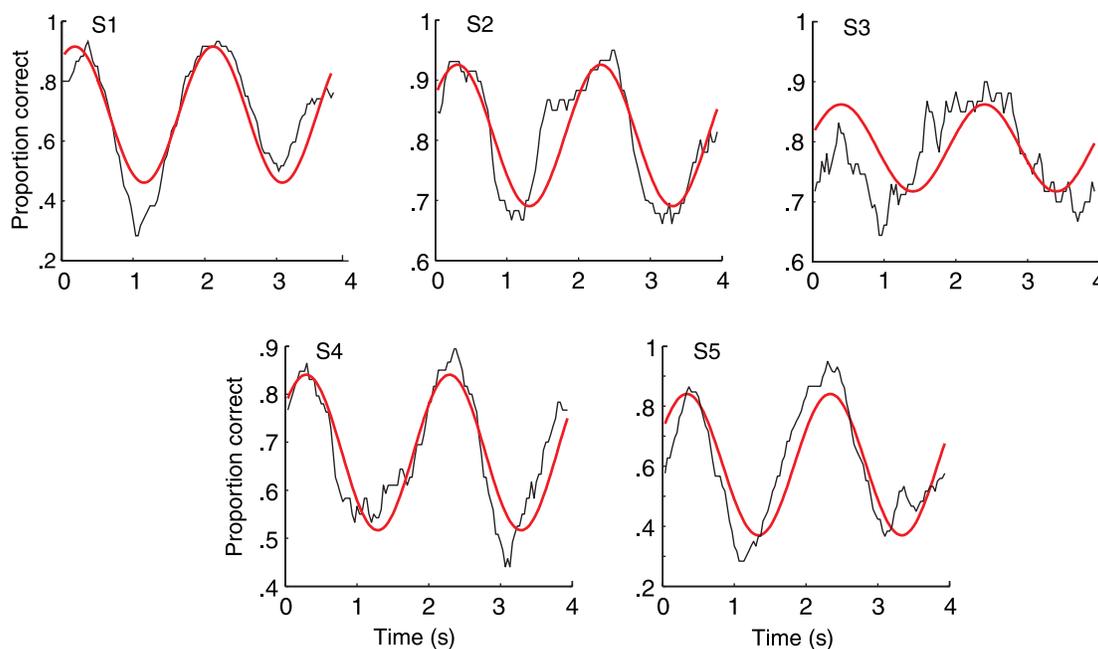
Figure 7. Individual subject performance in the jumping jack Bubbles experiment. Observer performance is plotted as a function of time (118 total frames at 30 Hz). Dark lines represent observer accuracy, and red lines are the least-squares sinusoidal fit to the data.

some aspects of the body, such as the shoulders, head, and hips, never experience instances of local opponent motion and therefore eliminating one of those dots does not change the opponent cues. We reason that these spatial Bubbles should reveal the relative importance of individual dots for discrimination performance.

Methods

Six unpaid participants were recruited for this experiment. As in Experiment 1, observers discriminated the facing direction of short (667 ms) point-light walker sequences embedded within a 3-s trial. Subjects were shown either the most diagnostic or least diagnostic intervals, as determined from Experiment 1. On each trial, a single, randomly selected dot was eliminated from the target figure. All other parameters were the same. Subjects completed a total of 600 trials or 300 trials per condition.

Proportion correct for the diagnostic and non-diagnostic trials were sorted based on dot omitted, and performance was collapsed across the extremities (ankles, knees, wrists, and elbows) and the non-opponent body parts (head, shoulders, and hips).

Results

On average, observers required about 80 noise dots ($SD = 17$) to maintain threshold discrimination performance. As expected observers had consistently worse

performance when viewing the non-diagnostic interval compared to the diagnostic interval (Figure 8). A repeated measures ANOVA revealed a significant main effect of interval ($p < .007$), as well as a significant interaction effect between interval and dots omitted ($p < .02$). A paired-samples t test revealed a significant difference between the body and the extremity dots during the diagnostic interval ($p < .0001$). Subjects are worse at discriminating the walker
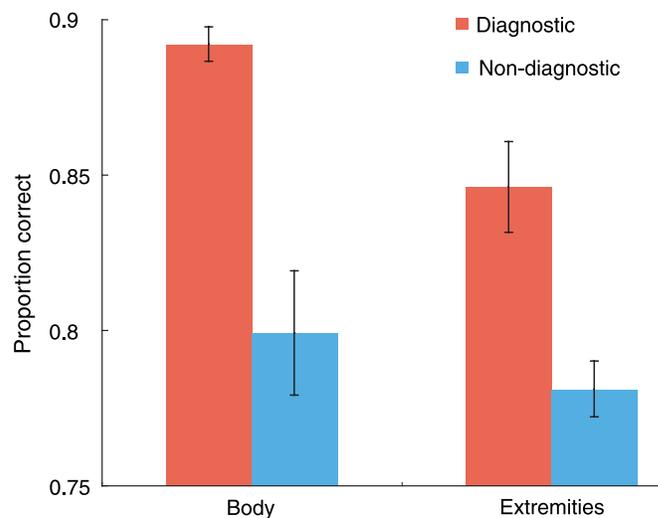

Figure 8. Results of the spatial Bubbles experiment. Mean accuracy ($n = 6$) for the diagnostic (red) and non-diagnostic (blue) intervals when body dots (head, shoulders and hips) or extremity dots (arms and legs) are omitted. Error bars indicate ± one standard error of the mean.

facing direction when the dots on the extremities were omitted during the diagnostic time interval, a manipulation that eliminates local instances of opponent motion. Omitting individual dots during the non-diagnostic intervals had no impact on performance. These results are consistent with the hypothesis that dynamic opponent motion cues are critical for detecting point-light biological motion.

Discussion

We have adapted the Bubbles technique from face and object perception to identify those features critical for the perception of point-light biological motion. This approach is not entirely unlike the temporal classification images generated by Lu and Liu (2006) for point-light biological motion animations. Their results suggest global analysis of body configuration for a complete walker gait cycle. These experiments document a number of new findings. First, perceptual sensitivity fluctuates over the course of a single point-light action. That is, certain moments in the action are more perceptually salient than others, which we argue correspond to those instances in which critical features are most apparent.

Second, we find that body posture is unlikely to play a critical role in discriminating point-light animations. This is in contrast to recent studies suggesting biological motion to be resolved via a form-based, template-matching algorithm. The form-based approach to solving point-light biological motion perception is inspired, at least in part, by studies demonstrating intact biological motion perception in animations with degraded local image motion (Beintema & Lappe, 2002). Point-light animations constructed with limited lifetime dots and dots jittered between the joints are readily recognized as biological motion. However, given the extended temporal summation window for biological motion (Neri et al., 1998), it is likely that some residual motion information remains available in these specially constructed displays (Casile & Giese, 2005).

Our Bubbles solution finds that the most perceptually salient moments are those in which the body posture is the least apparent. The most diagnostic moments in point-light animations are when the arms and the legs are closest and the body most vertically aligned. It is in the least diagnostic moments, when the arms and the legs are furthest apart in the gait cycle, that posture cues are the most apparent. It is thus unlikely that that static templates of body shape are critical cues in biological motion perception.

However, to test this hypothesis more directly, we measured discrimination accuracy for single frames. We found that when dynamic cues are not available, subjects better discriminate the frames with the most apparent body structure. In contrast, observers are least sensitive to the stationary images depicting narrow body structure (arms and legs crossing) that were the most diagnostic for the dynamic animations. Therefore, the strategies that are most effective for stationary point-light perception are the least effective for dynamic biological motion perception.

We conclude that dynamic cues are driving perceptual sensitivity to point-light biological motion. Observers most accurately discriminate biological motion when the relative distance between extremities is lowest, corresponding to local instances of opponent motion. This occurs when the legs cross during locomotion or the feet come together in a jumping jack. Likewise, eliminating a single dot from the extremities during the diagnostic intervals has the most detrimental impact on performance. The opposing motion of the joints appears to be a highly informative cue for detection and discrimination of point-light animations.

Horizontal opponent motion has been identified in earlier studies as a candidate critical feature for biological motion perception. Casile and Giese (2005) analyzed the form and the motion features in point-light and fully-illuminated versions of human actions. They found that only the motion features were shared by the two depictions, with the horizontally aligned opponent motion features dominating the motion domain. Moreover, they found that shuffling the overall spatial positions of these critical motion features does not eliminate the perception of biological motion. Observers spontaneously identify the animations as biological despite the impoverished form information.

Note that our results do not distinguish between models that propose biological motion to be subserved by specialized computations or by more general, flexible machinery. Relative motion is perceptually computed for a number of complex motion patterns, such as rotating wheels and bouncing balls (e.g., Johansson, 1974). It has been suggested that highly familiar complex patterns such as this could make up a "vocabulary" of sorts, for which we may develop dynamic templates (Cavanagh, Labianca, & Thornton, 2001). Because we so frequently encounter actions in our daily visual experiences, we are all highly trained experts on the range of familiar body movements. Biological motion would therefore be an ideal candidate for these putative dynamic templates.

If such templates did exist, then the likely neural instantiation would be the superior temporal sulcus (STS). It is within the STS that researchers have identified single neurons tuned to visual perception of specific body actions (Jellema, Maassen, & Perrett, 2004; Oram & Perrett, 1994). Neuroimaging studies in humans have identified the STS as having neural signals that highly correlate with the perception of bodies and body movements (for review, see Allison, Puce, & McCarthy, 2000). The STS is more strongly activated by point-light biological motion than by motion-matched controls, motion-defined objects, or complex articulating non-biological motion (Bonda, Petrides, Ostry, & Evans, 1996; Grossman

& Blake, 2002; Grossman et al., 2000; Pelphrey et al., 2003). The STS is activated during mental imagery of biological motion (Grossman & Blake, 2001) and by the implied biological motion in stationary silhouettes (Peuskens, Vanrie, Verfaillie, & Orban, 2005). Applying repetitive TMS over the STS in normal individuals impairs sensitivity to point-light biological motion perception (Grossman, 2005).

The human STS has been suggested as the integration site of motion and form computation in the construction of point-light biological motion (Giese & Poggio, 2003). It is the motion computations, however, that are the most critical for discriminating biological motion. In fact, creating a virtual "lesion" of the form pathway does not impact model performance. Our results similarly find that the motion computations are the most critical for biological motion perception. It remains to be seen whether these critical motion analyses are conducted within general motion-sensitive cortical brain areas (such as MT and MST), or within the biological selective STS region.

Conclusions

Perceptual sensitivity to biological motion waxes and wanes throughout point-light action cycles, depending on whether those moments with key features are more or less apparent. These changes in perceptual sensitivity are readily measured by the temporal "Bubbles" technique, modified to reveal key features in dynamic events. Point-light biological motion perception is most critically linked to relative motion of the joints, an inherently dynamic cue most apparent in the movements of the feet. These results are best predicted by computational models built from local motion features, but not by models that rely on stored templates of stationary body postures. Temporal Bubbles, as a general paradigm, is quite useful in revealing key features for dynamic event perception.

Acknowledgments

Corresponding author: Emily D. Grossman.
Email: grossman@uci.edu.
Address: 3151 Social Sciences Plaza, University of California Irvine, Irvine, CA 92697-5100, USA.

References

Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: Role of the STS region. *Trends in Cognitive Sciences, 4,* 267–278. [PubMed]

Beintema, J. A., & Lappe, M. (2002). Perception of biological motion without local image motion. *Proceedings of the National Academy of Science of the United States of America, 99,* 5661–5663. [PubMed] [Article]

Bonda, E., Petrides, M., Ostry, D., & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *Journal of Neuroscience, 16,* 3737–3744. [PubMed] [Article]

Bonnar, L., Gosselin, F., & Schyns, P. G. (2002). Understanding Dali's Slave market with the disappearing bust of Voltaire: A case study in the scale information driving perception. *Perception, 31,* 683–691. [PubMed]

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436. [PubMed]

Casile, A., & Giese, M. A. (2005). Critical features for the recognition of biological motion. *Journal of Vision, 5*(4):6, 348–360, http://journalofvision.org/5/4/6/, doi:10.1167/5.4.6. [PubMed] [Article]

Cavanagh, P., Labianca, A. T., & Thornton, I. M. (2001). Attention-based visual routines: Sprites. *Cognition, 80,* 47–60. [PubMed]

Cutting, J. E. (1978). Generation of synthetic male and female walkers through manipulation of a biomechanical invariant. *Perception, 7,* 393–405. [PubMed]

Cutting, J. E. (1981). Six tenets for event perception. *Cognition, 10,* 71–78. [PubMed]

Dittrich, W. H., Troscianko, T., Lea, S. E., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception, 25,* 727–738. [PubMed]

Gibson, B. M., Lazareva, O. F., Gosselin, F., Schyns, P. G., & Wasserman, E. A. (2007). Nonaccidental properties underlie shape recognition in Mammalian and nonmammalian vision. *Current Biology, 17,* 336–340. [PubMed] [Article]

Giese, M. A., & Poggio, T. (2003). Neural mechanisms for the recognition of biological movements. *Nature Reviews, Neuroscience, 4,* 179–192. [PubMed]

Gosselin, F., & Schyns, P. G. (2001). Bubbles: A technique to reveal the use of information in recognition tasks. *Vision Research, 41,* 2261–2271. [PubMed]

Grossman, E. D. (2005). Evidence for a network of brain areas involved in perception of biological motion. In

M. Grosjean, G. Knoblich, M. Shiffrar, & I. M. Thornton (Eds.), *The human body: Perception from the inside out* (pp. 361–384). Oxford: Oxford University Press.

Grossman, E. D., & Blake, R. (2001). Brain activity evoked by inverted and imagined biological motion. *Vision Research, 41,* 1475–1482. [PubMed]

Grossman, E. D., & Blake, R. (2002). Brain areas active during visual perception of biological motion. *Neuron, 35,* 1167–1175. [PubMed] [Article]

Grossman, E., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., et al. (2000). Brain areas involved in perception of biological motion. *Journal of Cognitive Neuroscience, 12,* 711–720. [PubMed]

Grossman, E. D., Blake, R., & Kim, C. Y. (2004). Learning to see biological motion: Brain activity parallels behavior. *Journal of Cognitive Neuroscience, 16,* 1169–1179. [PubMed]

Hiris, E., Humphrey, D., & Stout, A. (2005). Temporal properties in masking biological motion. *Perception & Psychophysics, 67,* 435–443. [PubMed] [Article]

Jellema, T., Maassen, G., & Perrett, D. I. (2004). Single cell integration of animate form, motion and location in the superior temporal cortex of the macaque monkey. *Cerebral Cortex, 14,* 781–790. [PubMed] [Article]

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics, 14,* 195–204.

Johansson, G. (1974). Vector analysis in visual perception of rolling motion. A quantitative approach. *Psychologische Forschung, 36,* 311–319. [PubMed]

Lange, J., Georg, K., & Lappe, M. (2006). Visual perception of biological motion by form: A template-matching analysis. *Journal of Vision, 6*(8):6, 836–849, http://journalofvision.org/6/8/6/, doi:10.1167/6.8.6. [PubMed] [Article]

Lu, H., & Lui, Z. (2006). Computing dynamic classification images from correlation maps. *Journal of Vision, 6*(4):12, 475–483, http://journalofvision.org/6/4/12/, doi:10.1167/6.4.12. [PubMed] [Article]

Mather, G., Radford, K., & West, S. (1992). Low-level visual processing of biological motion. *Proceedings of the Royal Society of London B: Biological Sciences, 249,* 149–155. [PubMed]

McCotter, M., Gosselin, F., Sowden, P., & Schyns, P. (2005). The use of visual information in natural scenes. *Visual Cognition, 12,* 938–953.

Nielsen, K. J., Logothetis, N. K., & Rainer, G. (2006). Discrimination strategies of humans and rhesus monkeys for complex visual displays. *Current Biology, 16,* 814–820. [PubMed] [Article]

Neri, P., Morrone, M. C., & Burr, D. C. (1998). Seeing biological motion. *Nature, 395,* 894–896. [PubMed]

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

Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10,* 437–442. [PubMed]

Pelphrey, K. A., Mitchell, T. V., McKeown, M. J., Goldstein, J., Allison, T., & McCarthy, G. (2003). Brain activity evoked by the perception of human walking: Controlling for meaningful coherent motion. *Journal of Neuroscience, 23,* 6819–6825. [PubMed] [Article]

Peuskens, H., Vanrie, J., Verfaillie, K., & Orban, G. A. (2005). Specificity of regions processing biological motion. *European Journal of Neuroscience, 21,* 2864–2875. [PubMed]

Pollick, F. E., Paterson, H. M., Bruderlin, A., & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition, 82,* B51–B61. [PubMed]

Schyns, P. G., & Gosselin, F. (2002). A natural bias for basic-level object categorizations [Abstract]. *Journal of Vision, 2*(7):407, 407a, http://journalofvision.org/2/7/407/, doi:10.11167/2.7.407.

Troje, N. F., & Westhoff, C. (2006). The inversion effect in biological motion perception: Evidence for a "life detector?" *Current Biology, 16,* 821–824. [PubMed] [Article]