

Joins and their relations as critical features in action discrimination: Evidence from a classification image method

Department of Psychology, University of California,
Los Angeles, Los Angeles, CA, USA
School of Psychological Sciences, Faculty of Medicine,
Nursing and Health Sciences, Monash University,
Clayton Campus, Victoria, Australia

Jeroen J. A. van Boxtel



Department of Psychology, University of California,
Los Angeles, Los Angeles, CA, USA
Department of Statistics, University of California,
Los Angeles, Los Angeles, CA, USA

Hongjing Lu



Classifying an action as a runner or a walker is a seemingly effortless process. However, it is difficult to determine which features are used with hypothesis-driven research, because biological motion stimuli generally consist of about a dozen joints, yielding an enormous number of potential relationships among them. Here, we develop a hypothesis-free approach based on a classification image method, using experimental data from relatively few trials (~1,000 trials per subject). Employing ambiguous actions morphed between a walker and a runner, we identified three types of features that play important roles in discriminating bipedal locomotion presented in a side view: (a) critical joint feature, supported by the finding that the similarity of the movements of feet and wrists to prototypical movements of these joints were most reliably used across all participants; (b) structural features, indicated by contributions from almost all other joints, potentially through a form-based analysis; and (c) relational features, revealed by statistical correlations between joint contributions, specifically relations between the two feet, and relations between the wrists/elbow and the hips. When the actions were inverted, only critical joint features remained to significantly influence discrimination responses. When actions were presented with continuous depth rotation, critical joint features and relational features associated strongly with responses. Using a double-pass paradigm, we estimated that the internal noise is about twice as large as the external noise, consistent with previous findings. Overall, our novel design revealed a rich set of critical features that are used in action discrimination. The visual system flexibly selects a subset of features depending on viewing conditions.

Introduction

The remarkable sensitivity of humans to biological motion has inspired a great deal of research directed at revealing critical visual features and robust representations used by the visual system to perceive biological movement (e.g., Cutting & Kozlowski, 1977; Johansson, 1973). This research has shown that humans can extract complex information from very impoverished biological motion displays (e.g., identity [Cutting & Kozlowski, 1977], emotion [Dittrich, Troscianko, Lea, & Morgan, 1996; Pollick, Paterson, Bruderlin, & Sanford, 2001], and type of action [Brown et al., 2005; Dittrich, 1993; Ma, Paterson, & Pollick, 2006; Norman, Payton, Long, & Hawkes, 2004; van Boxtel & Lu, 2011]).

Research in past decades suggests that the visual system employs two types of processing to accomplish this feat (Thornton, Pinto, & Shiffrar, 1998). The first type of processing is based on relational characteristics of multiple joints, which provide structural information in biological motion. For example, spatially scrambling or inverting an action animation makes biological motion harder to recognize or detect (Bertenthal & Pinto, 1994; Dittrich, 1993; Mark Williams, Huys, Canal-Bruland, & Hagemann, 2009; Pavlova & Sokolov, 2000; Proffitt & Bertenthal, 1990; Sumi, 1984), whereas making local motion information less informative (but not completely uninformative; Casile & Giese, 2005) does not hinder action perception (Beintema & Lappe, 2002). Others have shown that changing

Citation: van Boxtel, J. J. A., & Lu, H. (2015). Joins and their relations as critical features in action discrimination: Evidence from a classification image method. *Journal of Vision*, 15(1):20, 1–17, <http://www.journalofvision.org/content/15/1/20>, doi:10.1167/15.1.20.

limb configurations decreases performance in recognizing biological motion (Neri, 2009b; Pinto & Shiffrar, 1999), suggesting that some holistic information is provided by the correlated movements of the joints within a single limb.

The second type of processing is sensitive to information provided by the movements of some particular joints. For example, the feet have been shown to be very important in identifying the walking direction of a walker (Mather, Radford, & West, 1992; Thurman, Giese, & Grossman, 2010; Thurman & Grossman, 2008; Troje & Westhoff, 2006). More specifically, certain characteristics of the trajectories followed by the feet are critical in this type of discrimination task (Saunders, Suchan, & Troje, 2009), possibly more related to the velocity profile than to the exact shape of the trajectory (Hill & Pollick, 2000; Thurman & Lu, 2013a). Other work suggests that a potent cue in detecting walker stimuli is the presence of opponent motion signals (e.g., Casile & Giese, 2005; Thurman & Grossman, 2008). In ordinary circumstances, opponent motion is produced by the counter-phase motion of the limbs in natural human movements. However, even when opponent motion is introduced as part of an artificial stimulus, it can give the impression of biological motion (Casile & Giese, 2005). The dependency of biological motion perception on the two types of processing is further evidenced by a recent finding that the vertical position of the joint trajectories within the global layout of an action has a strong influence on the discrimination of walking direction (Hirai, Chang, Saunders, & Troje, 2011).

Given the involvement of different types of processing in biological motion perception, it is reasonable to expect that multiple sets of features play important roles in determining people's judgments when observing a particular biological motion stimulus. How can we determine in more detail which critical features are used by the visual system in recognizing biological motion? Typical approaches examine whether changes in certain stimulus parameters affect performance in perceiving biological motion, as measured by accuracy or response time. Measuring performance change as a function of certain stimulus parameters can indeed provide important evidence about whether those parameters affect the perception of biological motion. However, overall performance as measured by averaging responses across many trials is not sufficient to identify which aspects of the visual stimuli lead to a particular response on any individual trial.

In addition to these measurement-related issues, most studies in the literature have employed a hypothesis-driven approach. Although this approach can provide direct evidence to confirm or reject some specific hypothesis, in the case of biological motion perception, this general approach suffers from the

classical “curse of dimensionality”: biological motion stimuli have a large action space due to the many degrees of freedom in joint movements and the articulated body structure. It is therefore very difficult to test all possible parameter combinations or to make specific predictions. Hence, a data-driven approach may provide a complementary means as an association detector to extract the linkage between the presence of critical features in a complex input stimulus and a categorical output response. With the data-driven approach, the researcher does not limit the possible outcomes by designing an experiment such that it tests one specific hypothesis. Instead, the experiment allows the possibility for discovering many potentially interesting results.

A useful approach for this type of hypothesis-free experimentation is the classification image (CI) method (Ahumada Jr., 2002; Ahumada & Lovell, 1971; Beard & Ahumada, 1998; Eckstein & Ahumada, 2002; Neri & Levi, 2006). Although the internal representation of biological motion is not directly observable, it can be estimated by measuring the influence of certain characteristics of input stimuli on observers' responses. Specifically, CI methods examine trial-by-trial performance variations attributable to known amounts of noise added to stimuli, by computing the association between the noise and the observer's response (Ahumada Jr., 2002; Ahumada & Lovell, 1971; Gold, Murray, Bennett, & Sekuler, 2000; Victor, 2005).

In biological motion research, the CI approach has been used in several studies (H. Lu & Liu, 2006; Thurman et al., 2010; Thurman & Grossman, 2008; Thurman & Lu, 2013b). Results in some of these studies suggested that all joints contribute to detection of biological motion in luminance noise (H. Lu & Liu, 2006). Other studies used more refined judgments (e.g., walking direction discrimination) to reveal that the feet and wrist joints contribute most (Thurman et al., 2010; Thurman & Grossman, 2008; Thurman & Lu, 2013b). In all of these studies, as in similar approaches applied to face perception (e.g., Dupuis-Roy, Fortin, Fiset, & Gosselin, 2009; Kontsevich & Tyler, 2004), researchers have used luminance (pixel) manipulations to derive dynamic CIs based on >10,000 trials from one human participant to produce interpretable results. This inefficiency is largely due to the high dimensionality of added noise fields (i.e., adding pixel noise in image frames). Another limitation of previous CI studies is that the noise is added in a dimension (often luminance) along which biological motion perception is known to be invariant, which may weaken the successful application of CI methods in discovering some refined but important features involved in the perception of biological motion.

However, the efficiency of CI methods can be improved dramatically by selecting an appropriate

experimental design and developing advanced techniques for data analysis. One effective method of increasing the power of CIs is dimensionality reduction—selection of subspaces that are likely to contain a large signal-to-noise ratio (Eckstein & Ahumada, 2002; Ringach, Sapiro, & Shapley, 1997; Victor, 2005). For example, by reducing the dimensionality of positional noise, Li, Levi, and Klein (2004) successfully recovered CIs for position discrimination using just 750 trials. In a different approach, Ringach et al. (1997) incorporated a priori information in the design of the stimulus set to reduce the dimensionality of the input space, thereby effectively reconstructing veridical receptive fields of neurons using the reverse correlation between the input image sequence and the cell's spike train output. Another efficient method is to incorporate a priori knowledge about the target and dependency of input dimensions to allow pattern analysis in calculating CIs, rather than treating input dimensions independently as in the standard CI calculation (Neri, 2004; Victor, 2005).

The present article extends these methodological advances in the CI method to examine action discrimination, using a task in which observers view a morphed point-light action and classify it as a walker or a runner. To reduce the dimensionality of the analysis, we added noise to the morph weights of 13 joints in an action space, rather than introducing luminance noise in the image, as in previous studies. We took two prototypical actions, walking and running, and morphed each joint movement independently between these two actions. The resulting mean action is a 50%–50% morph to generate an ambiguous action between walking and running. On different trials, random Gaussian noise is added to introduce probabilistic morphing values for each joint independently. By adding noise to the joints in the action space instead of in the image space, we reduce the dimensionality of the CI analysis from thousands of dimensions (the number of pixels in the image) to about a dozen dimensions (the number of joints in a point-light actor). We thereby achieve interpretable CI results with many fewer trials (within hundreds of trials). Altogether, this method allows us to investigate, in a hypothesis-free manner, the importance of each of the joints and their relations in an action classification task.

The present article advances the CI method to reconstruct the critical features used by human observers during action discrimination tasks. The first experiment showed actions from a side view, a common viewpoint used in most biological motion research. In the second experiment, the actions were continuously rotated in depth over time. In addition to the viewpoint manipulation, a second independent factor (global body orientation) was introduced in both experiments. One condition consisted of intact

and upright actions; the other condition presented the inverted actions to reduce the involvement of structural processing in perception of biological motion. In the inverted condition, the difficulty of recognizing actions increases because of the reduced information about the human form. Therefore, this condition allows us to investigate whether critical features still play a role in biasing action discrimination judgment even with reduced familiarity to the inverted structure of the human actor.

Experiment 1

Methods

Participants

Twenty-two University of California, Los Angeles (UCLA), undergraduate students participated in the experiment for course credit. They were randomly assigned to one of two experimental conditions (upright and inverted). Two participants in the upright condition showed an extreme bias toward either walker or runner response (more than 90% of the trials were classified as walker or runner). These 2 participants were not further analyzed, resulting in 10 participants per condition (average age of 19.9 ± 1.8 , six men).

Stimuli

The motion capture data of walker/runner were obtained from the Carnegie Mellon motion capture database (<http://mocap.cs.cmu.edu>). The BioMotion-Toolbox (van Boxtel & Lu, 2013) was used in conjunction with the PsychToolbox (Brainard, 1997; Pelli, 1997), to display point-light stimuli, with 13 white dots (0.35°) representing the head, shoulders, hips, elbows, wrists, knees, and feet. The point-light actor was displayed from a sagittal (i.e., side) view, randomly selected to be leftward- or rightward-facing on each trial. The stimuli were 6° high and on average about $2\text{--}3^\circ$ wide, and they were displayed in the center of the screen on a gray background. The refresh rate was 75 Hz. The starting frame was randomly selected on each trial. We use orthogonal projections to display the actions.

The stimulus was an action morphed between a walker and a runner. The algorithm was adopted from the spatiotemporal morphing model developed by Giese and Poggio (2000). The morphed action was generated by linearly combining the movement trajectories of prototypical actions (walking and running) in three-dimensional space. One morph parameter, λ , controls the contributions of the individual prototypical actions to the linear combination of the algorithm,

so that morphed action = λ running + $(1 - \lambda)$ walking, with λ between 0 and 1. Thus, the value of the morph parameter controls the similarity of joint trajectories of a morphed action to the movements in prototypical actions. Other psychophysical studies used this algorithm to measure the generalization fields of action categories (Giese & Lappe, 2002) and to study action adaptation (van Boxtel & Lu, 2013). Although walking and running have different dynamics (Diedrich & Warren, 1995), previous research has shown that, for walking and running (and other bipedal locomotion), human observers can easily categorize the morphed actions, and they perceive them to be natural (Giese & Lappe, 2002).

In contrast to previous studies, which assigned the same morph value to all the joints, the present study created a morphed action, with each joint having an independently assigned morph value (see Figure 1A). Morph values were randomly sampled from a truncated Gaussian distribution with a mean of 0.5 and a standard deviation of 0.25, and the sampled values were bounded within 0 (walking) and 1 (running). Adding randomly sampled morph values to individual joints creates stimuli with added noise in a low dimensional space (i.e., 13 in our study) but still maintains sufficient trial-by-trial variations to derive CIs from participants' responses (see Li et al., 2004, for a similar approach in a position discrimination task).

As shown in Figure 1B, the experiment included two conditions: (a) upright point-light actors with intact body structure and (b) inverted point-light actors with intact body structure.

Procedure

The experiment included 1,200 trials, each lasting 1 s, containing exactly one walking/running cycle. The first 1,000 trials used stimuli that were generated according to the morphing methods described above. The last 200 trials were trial-by-trial repetitions of trials 800 to 1,000, which were used to assess intraindividual consistency, by calculating the proportion of identical responses on identical trials. This double-pass procedure allows the estimation of the internal noise inherent to the system (i.e., the brain) relative to the external noise added to the stimuli (Burgess & Colborne, 1988; Z. L. Lu & Doshier, 2008; Neri, 2009a).

After the stimulus disappeared, participants were asked to indicate whether the observed actor appeared to be a walker or a runner, using the left or right arrow keys on the keyboard. Before the experiment started, all observers (including those assigned to the inverted conditions) performed practice trials on upright actors until they indicated they felt comfortable with the task (generally 5–10 trials). They were then told to do the same task on the stimuli that were presented in the

experiment, regardless of the body orientation of the actor. The entire experiment took about half an hour.

Analysis

A logistic regression model was used to analyze the relationship between added morph values for individual joints and the participants' responses (i.e., walker or runner).¹ Given that the classification responses were binary variables, logistic regression analysis incorporates a nonlinear function to estimate a weighted combination of joint morph values in predicting the probability for a "walker" response:

$$P(\text{walker}) = \frac{1}{1 + e^{-(\beta_0 + \sum \beta_i \lambda_i)}}$$

where i indicates the joint index. With 1,000 trials of morph values used in stimulus generation and corresponding responses, the maximum likelihood estimation yielded two important assessments on (a) the importance of the joints reflected by the regression weights for joints (i.e., beta weights in the above logistic equation) and (b) the relations between joints reflected by the correlations between the beta weights (see Figure 2). The correlations between beta weights, obtained per subject, shed light on how multiple joints were used simultaneously and thus revealed important aspects of relational processing in action discrimination.

The estimated beta weights of identical joints on both sides of the body were averaged per subject (except the head joint, which lacks a mirror-positioned joint). This was done to increase sensitivity and because the stimulus is ambiguous as to what is the left and right side of the body. Similarly, correlations that were in mirror positions were also averaged. The obtained beta weights and correlation values were then used in statistical analyses performed across subjects. We performed one-sample two-tailed t tests, which were corrected for false discovery rate (Benjamini–Hochberg procedure) for the correlation analyses within each experimental condition.

Results

CIs for upright actors in a side view

Figure 3A shows the resulting CIs, including beta weights and their correlations, for an upright actor in a side view. The foot joints showed the most significant contribution as revealed by the largest beta weight, $t(9) = 9.58$, $p < 0.001$, Cohen's $d = 3.0$, indicating that the similarity of the observed motion trajectories of the feet to the prototypical movements played the most critical role in discriminating walking from running, consistent with previous findings (Mather et al., 1992; Thurman et

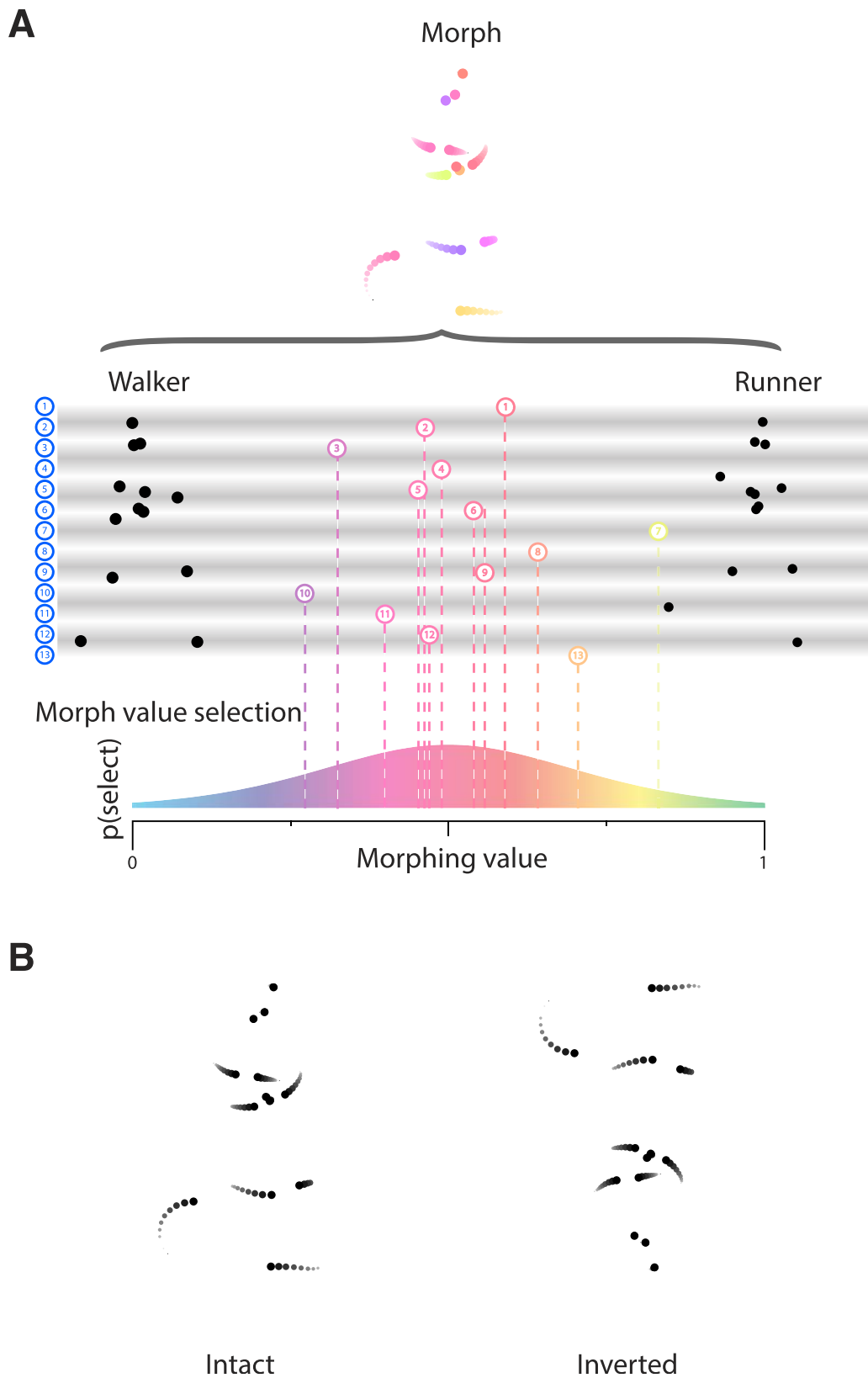


Figure 1. Stimulus design. (A) Selection of morphing values. For each joint, a morphing value is drawn independently from a truncated normal distribution, schematized at the bottom of Panel A. Each joint's movement will be a morphed trajectory between walking and running, based on the morphing weight drawn for that joint. The joints' movements will be played together as a single morphed



←

action (top). (B) Schematized stimuli in the intact and inverted conditions. To illustrate motion, the dot size is increased from frame to frame and the color is changed from white to black. This is only for illustrational purposes; in the experiments, the dots' appearance remained unchanged. Every fifth frame is drawn, for the first 50 frames.

al., 2010; Thurman & Grossman, 2008; Troje & Westhoff, 2006).

However, the feet were not the only joints showing significant contributions in distinguishing walkers from runners. In fact, most joints, except the hips and the knees, are depicted with red-framed circles in Figure 3A, because of the significantly positive beta weights associated with these joints, indicative of a contribution to the discrimination responses: head, $t(9) = 2.35$, $p = 0.04$, Cohen's $d = 0.7$; shoulder, $t(9) = 4.4$, $p = 0.002$, Cohen's $d = 1.4$; elbow, $t(9) = 2.3$, $p = 0.05$, Cohen's $d = 0.7$; wrist, $t(9) = 5.0$, $p < 0.001$, Cohen's $d = 1.6$. This abundance of significant beta weights for most joints suggests that, when viewing a walker or a runner from a side view, the brain employs a holistic structural processing. The coexistence of critical joint features (e.g., feet) and holistic structural

features (i.e., contributions of most joints were significant) provides converging evidence that the visual system identifies actions by using both local mechanisms based on the most discriminative movement features and global mechanisms based on holistic structure.

However, it is worthwhile to note that some of the significant contributions of joints obtained in the current study may be specific to classifying a walker from a runner. For example, the head was associated with a significant weight, showing a significant contribution to the response. This relationship is probably due to the fact that during running, the actor's head leans more forward than during walking. Such a cue from head movement or position may not be informative for, for example, judging forward/backward walking direction; thus, head movements are unlikely to be identified as critical features in that particular task.

The resulting CIs for the upright actors in a side view also signaled the presence of relational features through significant correlations between the beta weights of joints. There were two significantly positive correlations between beta weights, depicted as colored lines in Figure 3A. First, there was a significant correlation between the two foot weights, $t(9) = 4.3$, $p = 0.0019$, Cohen's $d = 1.4$, suggesting that observers use information from both feet simultaneously when performing the walking/running discrimination task. The discovery of the relational feature from the resulting CIs is consistent with previous findings about the importance of opponent motion signals in identifying actions involving bipedal locomotion (e.g., Casile & Giese, 2005; Thurman & Grossman, 2008). The other significant correlation was between the wrists and the opposite hip joint, $t(9) = 5.5$, $p = 0.0004$, Cohen's $d = 1.7$. This correlation suggests that observers assess a relational feature—the movement of the wrist relative to the hips—to determine whether the actor is walking or running (i.e., judging whether the wrist is low or high, respectively, with respect to the hips). This relational feature, to our knowledge, has not been reported and examined in the literature. This finding shows that humans employ a rich set of relational features, not solely constrained by limbs and body structure, to facilitate action discrimination.

In this experiment, we analyzed the last 400 (200 + 200) trials with the double-pass procedure to examine the consistency of subject responses. We found that

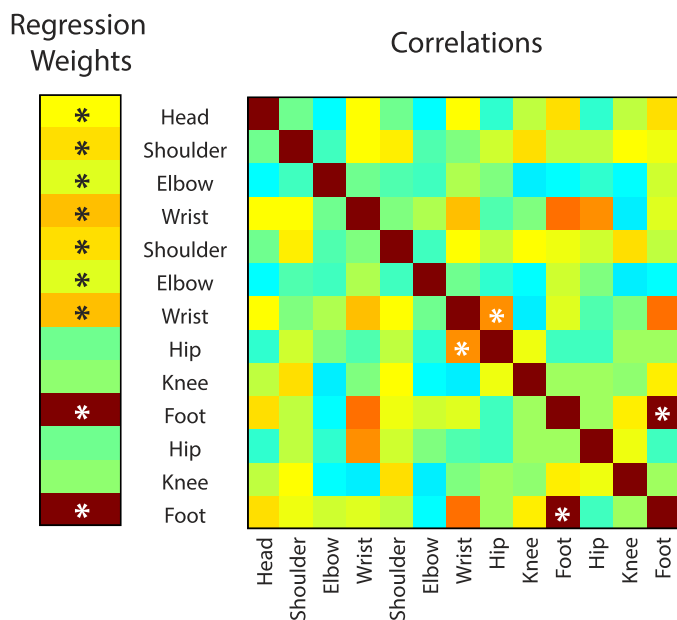


Figure 2. Logistic regression results. The logistic regression returns a vector with regression weights. These weights are representative of the importance of the joints to the classification task. In the remainder of the article, these weights are displayed on the joints of the walker (see e.g., Figures 3 and 4). The logistic regression also returns a correlation matrix, with correlations between all joints. These are displayed as lines between joints in the remainder of the article. This figure shows the results averaged over subjects for the upright side view condition. Stars indicate significance.

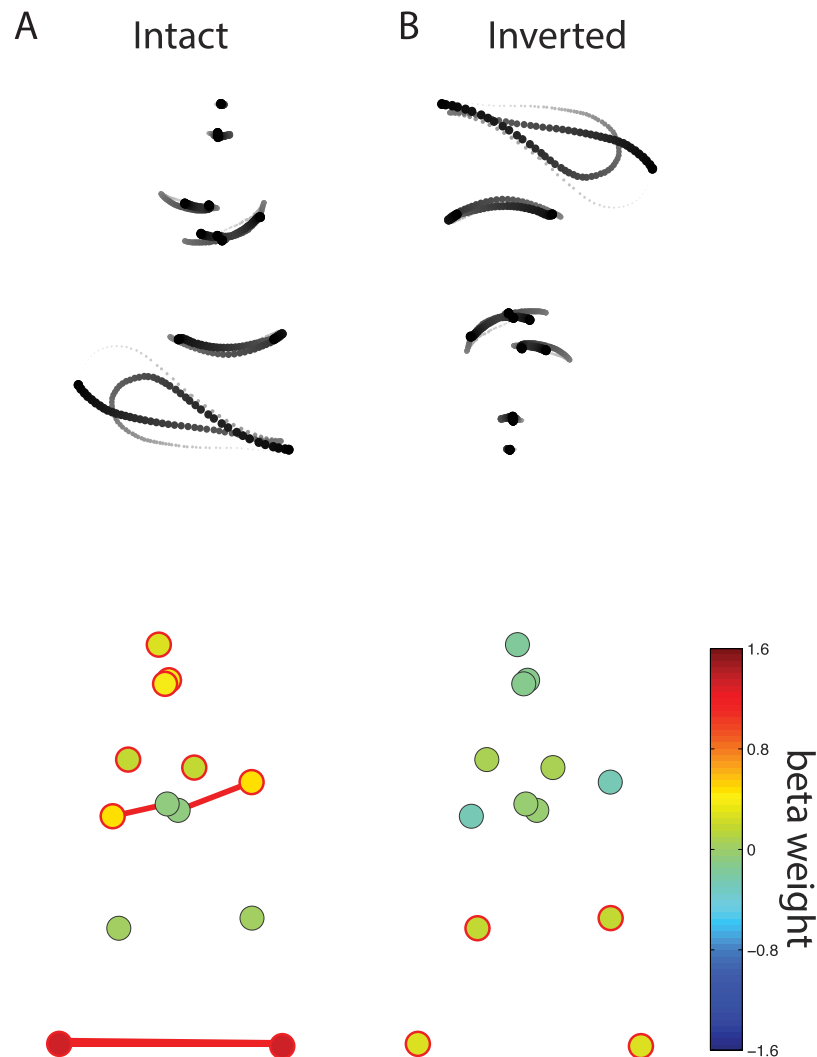


Figure 3. Classification data from the binomial logistic regression for Experiment 1. Biological motion actions were observed from a side view. The upper panels show the trajectories of a single trial. The bottom panels show the resulting classification images. The magnitude of the beta weights is indicated by different colors, with green being neutral and blue indicating negative weights, whereas red indicates positive weights. Significant deviations from zero are indicated by a red circle around the joint. Correlations are indicated by colored lines, connecting joints. Significantly positive correlations are colored red, and significantly negative correlations are colored blue. Only the upright condition showed significant correlations to indicate the involvement of relational processing. It also showed indications of a more holistic template-matching processing, as the whole upper body was significant. The inverted condition yielded critical joint features only in the legs.

intraindividual consistency was 0.56 ± 0.019 , which, though somewhat low, was significantly better than chance level, $t(9) = 3.3$, $p = 0.009$, Cohen's $d = 1.48$. This result suggests that individuals maintained a certain level of consistency in discriminating highly ambiguous actions, when confronted with the same stimuli at different times.

CIs for inverted actors in a side view

The same analyses as explained above were performed for the inverted condition. Compared with the upright condition, the resulting CIs (Figure 3B) revealed rather fewer joints with significant contribu-

tions to action discrimination. Only the feet and knees reached significance: feet, $t(9) = 3.0$, $p = 0.015$, Cohen's $d = 0.95$; knee joints, $t(9) = 2.3$, $p = 0.04$, Cohen's $d = 0.7$. This pattern of results echoes previous findings indicating that the feet are very important joints in a walking direction discrimination task, even in inverted conditions (Chang & Troje, 2009; Thurman & Lu, 2013b). However, the lack of significant contributions from other joints indicates the weak involvement of holistic processing when viewing inverted actions. Furthermore, the resulting CI did not reveal any relational features based on significant correlations between joints, confirming the reduction of structural

processing when actions are inverted. Intraindividual consistency was 0.63 ± 0.022 , showing a significant difference from random responses, $t(9) = 6.0$, $p = 0.0002$, Cohen's $d = 2.68$.

Experiment 2

Experiment 2 used the same stimulus generation and procedures as in Experiment 1, except that the actions continuously rotated in depth. With the inclusion of three-dimensional (3D) depth rotation, the individual dots' trajectories projected to the two-dimensional image plane are the result of a combined action-related joint movement and the 3D rotation. This complexity will disrupt the utility of simple opponent motion analysis strategies (Casile & Giese, 2005) or strategies based on some characteristic features of feet movement trajectories (Saunders et al., 2009). This decreased discriminability in the trajectories of local dot movements may encourage participants to exploit relational features when making discriminations between walkers and runners.

Methods

Twenty-two UCLA undergraduate students participated in Experiment 2 for course credit. These students had not participated in Experiment 1. They were randomly assigned to one of two experimental conditions (upright, inverted). Two participants in the inverted condition showed an extreme bias toward either walker or runner response (i.e., more than 90% of the trials were classified as walker or runner). The results of these two participants were not further analyzed, resulting in 10 participants per condition (average age of 20.0 ± 1.5 , eight men).

In this experiment, stimuli were generated and presented in the same manner as in Experiment 1, except that the presented actions rotated in depth. On each trial, the action was rotated in depth at a speed of $1^\circ/\text{frame}$ ($75^\circ/\text{s}$). Because we used an orthogonal projection, the rotation direction could be perceived as either clockwise or counterclockwise. The procedure and analysis were the same as used in Experiment 1.

Results

CIs for upright actors in rotating view

Consistent with the findings from the upright condition in the side view, the resulting CIs revealed that the feet contributed significantly to the discrimination judgment even when the actions were presented in rotating views, $t(9) = 4.36$, $p = 0.0018$, Cohen's $d =$

1.38, as shown in Figure 4A. However, the beta weight for the feet was smaller in the rotating view than in the side view in Experiment 1, $t(18) = -2.74$, $p = 0.014$, Cohen's $d = -1.22$. Again, the wrists were the second most important joints, $t(9) = 2.66$, $p = 0.026$, Cohen's $d = 0.84$. The beta weight for the wrist was also smaller in the rotating condition than in the side view, $t(18) = -3.85$, $p = 0.0012$, Cohen's $d = -1.72$. Even though the contributions from the feet and wrist joints are smaller in the rotation condition, revealing the consistent importance of these joints in both experiments shows the generalization of critical joint features to more complex viewing conditions.

As shown in Figure 4A, other joints failed to make significant contributions to judgments of action discrimination, suggesting a much weaker involvement of holistic features in structural processing when actions were under 3D rotation. The inclusion of the depth rotation in this experiment likely weakened form-based mechanisms based on template matching over frames (Lange & Lappe, 2006). Hence, the failure to find any structural features for rotating actions in 3D may be due to the reduced activity of snapshot (Vangeneugden, Pollick, & Vogels, 2009) and viewpoint-dependent action detectors (Perrett et al., 1985; Vangeneugden et al., 2011).

For relational features, the resulting CIs revealed a negative correlation between the elbow and the hips on the same side of the body, $t(9) = -6.1$, $p = 0.00018$, Cohen's $d = -1.92$. We interpret the negative correlation as an indicator that subjects selectively pay attention to either elbow joint or the hip joint at a given time. With the depth rotation, it is likely that participants used hip movements to infer a 3D body facing direction, so that a 3D body layout can subsequently be established as the reference, and then compare the elbow movements to the body layout reference to determine whether the actor is running (sharp angle between elbow, wrist, and shoulder joints) or walking (obtuse angle between these joints). Note that the movements of elbows themselves were not critical joint features, as indicated by the lack of significant beta weights. However, the relative relations of elbows to hip joints did play an important role in determining participants' responses. Intraindividual consistency was significantly above chance level, 0.65 ± 0.046 , $t(9) = 3.46$, $p = 0.007$, Cohen's $d = 1.5$. To our knowledge, this is the first evidence supporting the important role of a relational feature based on the movements of elbows and hips in recognizing actions in 3D with rotating view.

CIs for inverted actors in rotating view

The resulting CIs (Figure 4B) revealed two critical joint features in discriminating inverted actors with rotating view: the feet, $t(9) = 2.78$, $p = 0.021$, Cohen's $d =$

$= 0.88$, and the knees, $t(9) = 2.72$, $p = 0.024$, Cohen's $d = 0.86$. These features are consistent with the results from the inverted condition with the side view in Experiment 1, suggesting the robustness of these critical features when observing inverted actions across different viewing conditions. The analysis did not yield any significant correlations between joints, confirming the reduced structural processing for inverted actions. Intraindividual consistency was only slightly higher than chance: 0.55 ± 0.02 , $t(9) = 2.65$, $p = 0.026$, Cohen's $d = 1.19$.

Additional analysis

Principal component analysis

We also performed principal component analyses (PCAs) on the data. For each individual subject, we calculated the average morph weights from all trials reported as a walker and the average morph weights from the trials reported as a runner. These average weights were put into one vector with the length of 13 (i.e., the number of joints) for each action type individually. From 10 participants, this operation resulted, per condition, in 10 walker vectors (i.e., 1 vector per subject) and 10 runner vectors. These vectors were then concatenated into one matrix in the size of 20×7 (10 walker averages + 10 runner averages times 7 joint pairs [head, shoulders, etc.]) to provide the input for the PCA.

As a baseline comparison, we calculated the expected contribution of each component based on chance, according to the broken stick model (Frontier, 1976). The broken stick model randomly divides a line of unit length (representing total variance) repeatedly into a number of segments equal to the number of PCA components and sorts these in descending order. The average lengths of the ordered segments over repeated permutations represent the eigenvalues expected by chance. It is only informative to interpret those PCA components that have a larger contribution than the broken stick model (i.e., a larger contribution than expected by chance).

For Experiment 1 in a side view, the PCA overall confirmed the main results revealed with the CIs (see Figure 5). In both the upright and the inverted side-view condition, only the first of the principle components (blue bars) explained more of the variability among participants than expected by chance (red line). Furthermore, the contributions of each joint can be quantified as the absolute value of loading scores to the component given by the PCA. For the upright actors in the side view, the PCA revealed that the first principle component loaded most heavily on the feet. Meanwhile, the loadings from joints of the upper body also contributed considerably, consistent with the analysis on the CIs in

above sections. For the inverted actors, the first principle component loaded primarily on the feet and the head, with a gradual transition over the entire length of the body. A similar, but less apparent, trend was observed in the CI analysis.

For Experiment 2 with the rotating view, the PCA (Figure 5) showed that, only in the upright condition, the first principle component accounted for more variability than expected by chance, and it loaded primarily on the feet. On the contrary, in the inverted condition, none of the principle components explained the variability better than expected by chance.

Bias and consistency values: Internal noise estimates

We analyzed the consistency values through a double-pass procedure (Burgess & Colborne, 1988; Z. L. Lu & Doshier, 2008; Neri, 2009a). Identical trials, with exactly the same added noise, were presented to the subject twice, and we recorded whether the answer was identical in the two passes. From the double-pass design, we can measure consistency as the proportion of identical trials in which the participant gave an identical answer and bias as the proportion of trials that were given one of both responses (e.g., the overall proportion of trials that was reported as a walker). We took the largest of the two biases as our bias measure: $\text{bias} = \max(\text{P}(\text{report_walker}), \text{P}(\text{report_runner}))$.

We found that several observers had strong biases toward either runner or walker. Such strong biases inflate consistency values. For example, a participant providing the same response for all trials would achieve 100% consistency between the two double-pass blocks. Hence, there is a clear monotonic relationship between the bias and the consistency values. This theoretical relationship between bias and consistency is captured in Figure 6. A gray area in Figure 6 depicts the theoretically impossible range of consistencies given a certain bias. A gray curve in Figure 6A and B is used to show an average consistency when a participant responds with a certain bias and is otherwise influenced only by internal fluctuations in the system. In this case, the response is based on internal noise and not on the signal. In other words, the consistency is based solely on the bias and not on the use of information present in the stimulus.

Nevertheless, notwithstanding the greater consistency with the increase of bias, most observations still lie above the curve derived from random responses (trending with a nonparametric binomial test over all data; $p = 0.08$) suggesting that participants responded in a consistent way despite the ambiguity in the morphed actions.

When performing double-pass experiments, studies in the literature generally plot consistency (often called agreement) as a function of accuracy to examine how

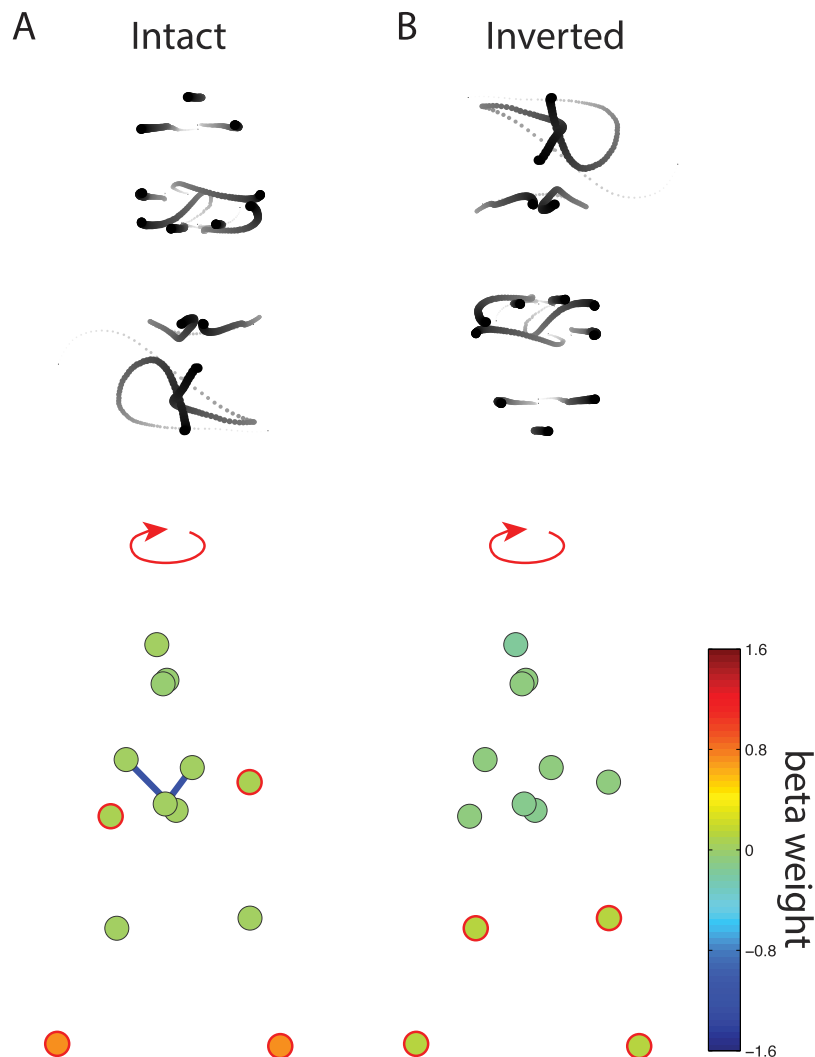


Figure 4. Classification data from the binomial logistic regression for Experiment 2. Biological motion actions were rotating in depth. The upper panels show the trajectories of a single trial. The bottom panels show the results. As in Figure 2, the magnitude of the beta weights is indicated by different colors, with green being neutral and blue indicating negative weights, whereas red indicates positive weights. Significant deviations from zero are indicated by a red circle around the joint. Correlations are indicated by colored lines, connecting joints. Significantly positive correlations are colored red, and significantly negative correlations are colored blue. Only the upright condition showed significant correlations.

consistency and accuracy relate to each other and to estimate external and internal noise values (as well as other assumptions; Burgess & Colborne, 1988; Z. L. Lu & Doshier, 2008; Neri, 2009a). However, in our study, we presented ambiguous actions between walker and runner with a 50–50 morphing value; the accuracy measure was not manipulated. However, we can plot the relationship between bias and consistency (as shown in Figure 6). We derived the relationship between bias and consistency for various ratios of external and internal noise through computer simulations. The simulations assume an internal noise that is normally distributed, with a mean of x and a standard deviation of 1. To obtain a range of biases, x is varied. A value for this internal noise is randomly sampled on

each trial. The external noise is also normally distributed, with a mean of 0 and a standard deviation as a certain fraction of the deviation used for the internal noise. This external noise is the same on both passes in the double-pass simulations. The total signal in a trial is the addition of the internal and external noise. Consistent trials are those double-pass trials that are both larger than zero or both smaller than zero. Bias is calculated as defined above. When the results in Experiments 1 and 2 are replotted in Figure 6C and are compared with the curves relating bias to consistency at various external/internal noise ratios, we find that our data cluster close to a ratio of external/internal noise equal to 0.5. This value is comparable to, but lower than, earlier estimates for other tasks and stimuli (Neri,

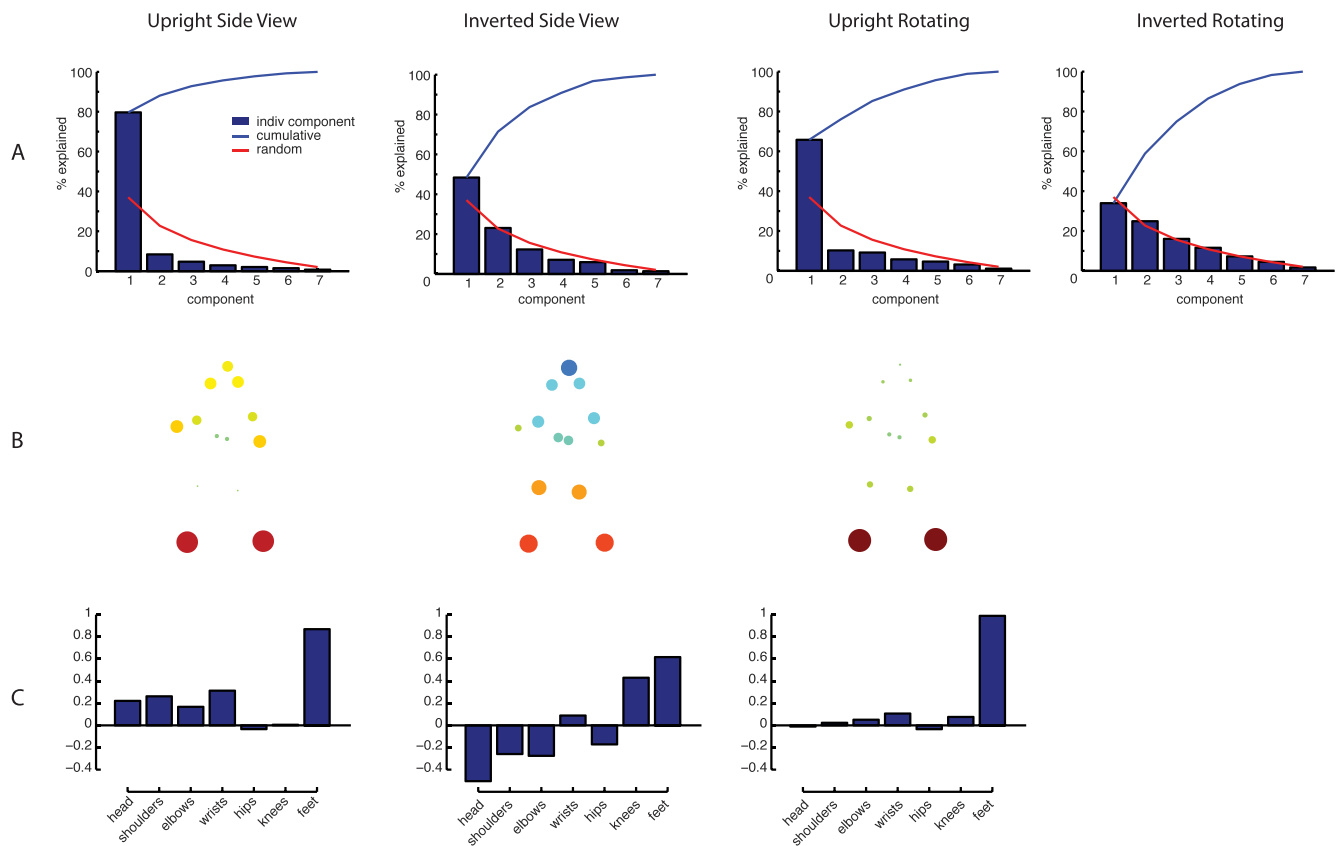


Figure 5. Results of the principal component analysis. Plots in Panel A show the individual contribution of the different principal components to the overall variance of beta weights across subjects (bars) and the cumulative contribution of all components up to the component mentioned on the x-axis (blue line). The red line shows the contribution of each component expected by chance (broken stick model). The result shows that only the first component contributes more than expected by chance in the two upright conditions and the inverted side-view condition. Figures in Panel B show the loadings of the first component in the layout of a point-light actor, in which the red indicates positive loading values and the blue negative values. In Panel C, the same data used in Panel B are shown but as a bar chart.

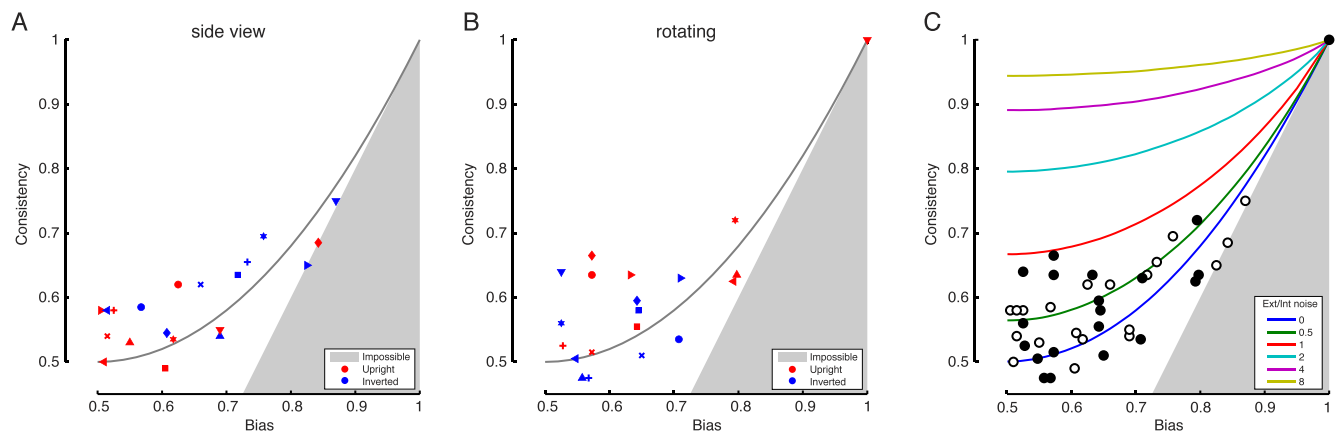


Figure 6. Results from the double-pass trials. (A) Bias and consistency values are plotted for individual participants (different symbols), for the upright (red) and the inverted (blue) conditions, in Experiment 1 (side view). (B) Results from Experiment 2 in rotating view, from different participants from Experiment 1. (C) The same data were replotted for Experiment 1 (open symbols) and Experiment 2 (filled symbols) overlaid on the curves showing the relationships between bias and consistency dependent on different ratios of external over internal noise.

2010), which were closer to 1, with the average value around 0.8. Using $1/(1 + x^2)$ (Neri, 2010), with x being the ratio of internal/external noise (2 in our case), our noise estimate translates into an efficiency of about 0.2, meaning that subjects use about 20% of the statistical information available in the stimulus (Barlow, 1978; Neri, 2010).

Discussion

Human are exceedingly sensitive to structured motion signals, both animate and inanimate (Hiris, 2007). Using a hypothesis-free reverse correlation technique, we aimed to unfold the critical features used in action discrimination. We show that humans have an efficiency of 20% in extracting information from our biological motion stimuli. This information, we show, is carried by three types of features, which play important roles in discriminating bipedal locomotion presented in a side view: (a) critical joint features supported by the finding that the similarity of the movements of feet and wrists to prototypical movements of these joints were most reliably used across all participants; (b) structural features reflected by small but significant contributions from movements of almost all other joints, which potentially underlie a form-based analysis; and (c) relational features, revealed as statistical correlations between joint contributions, including relations between the two feet, and between the wrists/elbow and the hips. When the actions were inverted, only critical joint features showed a significant influence in the discrimination judgment; structural and relational features were not revealed to show any contributions. When the actions were presented with continuous depth rotation, critical joint features and relational features exhibited strong associations with the responses, but there was an absence of structural features. Overall, our novel design shows that a rich set of critical features is used for the perception of biological motion, and the visual system flexibly selects a subset of features depending on viewing conditions to facilitate action discrimination. Below we will discuss the importance of these findings and clarify the connections with the literature.

Efficiency

We found that humans have an efficiency of about 20%, which would mean that they use 20% of the statistical information available in the stimulus in our action discrimination task. This value is on the lower end of reported efficiencies for many low-level tasks

(Neri, 2010). These earlier estimates of internal noise were based on stimuli that were presumably processed in earlier stages of the visual system (luminance, contrasts) than our stimuli (biological motion). The higher-level processing may be one possible source for the high internal noise and low efficiency in action discrimination. The finding of lower efficiencies of about 8% in a 3D shape recognition task (Tjan, Braje, Legge, & Kersten, 1995) supports this interpretation.

The efficiency value of $\sim 20\%$ in our study of action discrimination is close to the efficiency found in gender and affect recognition ($\sim 30\%$) from biological motion (Pollick, Lestou, Ryu, & Cho, 2002). Human efficiency in action discrimination, however, is much higher than the efficiency in identifying the walking direction when biological motion stimuli are embedded in luminance noise ($< 1\%$; Gold, Tadin, Cook, & Blake, 2008). This difference suggests that the human visual system is more efficient in processing the information about movement trajectories of joints in biological motion than processing the luminance information by matching to posture templates.

Critical joint features for action discrimination

Feet

Previous studies have found that the feet are important for judging the walking directions of a point-light actor (Mather et al., 1992; Thurman et al., 2010; Thurman & Grossman, 2008; Troje & Westhoff, 2006). The reason for the importance of the feet is probably twofold. First, the feet, when most extended, play an essential role in providing a recognizable skeleton shape for the body from a side view, which can be used for detecting the facing direction using the key-frame template (Thurman & Grossman, 2008). Second, for actions with bipedal locomotion (e.g., walking, running), the movements from the two feet generate an opponent motion signal when they cross midcycle. Such opponent motion can be readily extracted by motion-sensitive neurons in the middle temporal area (Heeger, Boynton, Demb, Seidemann, & Newsome, 1999) to promptly signal potential human movements and elicit the impression of biological motion (Casile & Giese, 2005; Thurman & Grossman, 2008).

In the present study—involving the discrimination of different actions, and not walking directions—we found that the feet were indeed most consistently used by the observers, especially in the upright conditions, and to a lesser extent in the inverted condition. The fact that the hypothesis-free approach reveals the same critical features as previous research confirms the value of this method.

Interestingly, we also find a decrease in the use of the feet (and the wrists) when the action is rotated in depth as compared to when the action is shown from a side view. We explain this by noting that trajectories of individual joints and opponent motion signals are much less informative in this condition, because the rotation of the actor itself causes extrinsic motion signals not due to the actor's movements and spurious opponent motion signals that are not informative about the action portrayed.

Wrist, elbow, and knees

Importantly, the hypothesis-free approach enabled us to determine the contribution of the other joints in the action discrimination task. In addition to the feet, the wrists appeared as the second most important joints in the two upright conditions. This finding is consistent with previous findings that the movements of wrists have also been identified as critical features in judging walking direction of an actor, (Mather et al., 1992; Thurman et al., 2010) and animacy ratings (Thurman & Lu, 2013b) but less so in detection tasks (Pinto & Shiffrar, 1999). Hence, the contribution of the wrists varies depending on the nature of different tasks.

In our action discrimination task, the wrists may have contributed in at least two different ways to the performance. First, they are, after the feet, the joints that produce the largest excursion and the largest opponent motion signals, which provide strong cues for form-based and motion-based analyses. Second, the position of the wrist determines in large part the arm angle subtended by the shoulder, elbow, and wrist (this is so because the elbow and shoulder do not differ as much as the wrist between walker and runner). The arm angle itself can be a very informative cue for the discrimination tasks, with sharp angles indicating running and obtuse angles (stretched arms) indicating walking.

Consistent with previous literature (Mather et al., 1992; Thurman et al., 2010; Thurman & Lu, 2013b), we found that the knees did not contribute to the discrimination task in upright conditions. The low importance may be somewhat surprising, because the knees (as well as the elbows, which contribute little as well) exhibit significant excursions in 3D space and could potentially be used by the observer as a discriminative feature. One reason why these joints may be unused in the current task is that there is enough information in the wrists and feet, obviating the need for the observer to use the information from the knees and the elbows, even though they do contain discriminative information. However, it seems unlikely that the knees contain much information in ecologically valid upright body orientations, because it was previously reported (Mather et al., 1992) that

when the wrist and feet are removed from the display, the subjects' performance on a direction discrimination task dropped to chance levels. This finding indicates that observers really do not use the information from the knees and elbows, even when the wrists and feet are removed. However, it should be noted that the task in the study by Mather et al. (1992) was different from our task (i.e., a walking direction discrimination task and not an action discrimination task).

Structural processing

A striking observation in our results is that, for the upright actions in the side view, almost the entire set of joints plays a significant role in affecting action discrimination judgments, including all the joints of the upper body. Clearly, in the side view, human observers use a wide array of cues originating from the entire stimulus, and not just from the feet, suggesting a more holistic processing involved in action discrimination. This finding accords well with a previous analysis of emotional gait patterns that indicated that most if not all joints contain, in principle, discriminative information about inferring emotion from observed actions (Roether, Omlor, Christensen, & Giese, 2009). In another study (Mark Williams et al., 2009) it was shown that inferring arm movement in tennis also depended on nonarm joints.

The PCA provides perhaps the most direct way of establishing the contribution of structural processes. Overall, the PCA provided data very consistent with those from the CI analysis. The PCA revealed that the upright side view condition showed significant structural processing, whereas the other conditions provided much less support for the influence of joints beyond the feet in the classification process.

The fact that observers use such a wide range of joints in our experiments, but only in the side view condition, could be an indication of the use of a template-matching procedure in a familiar and common viewpoint from the sagittal viewing angles (Lange & Lappe, 2006).

Relational features

Previous research has repeatedly shown that biological motion perception depends not only local processing but also partly on holistic processing, taking into account the spatiotemporal relationships between joints (e.g., Thornton et al., 1998). Spatially scrambling or inverting an action animation decreases recognition and detection performance (Bertenthal & Pinto, 1994; Dittrich, 1993; Pavlova & Sokolov, 2000; Proffitt & Bertenthal, 1990; Sumi, 1984). In other research, it was shown that correct limb configurations are important

for biological motion perception (Neri, 2009b; Pinto & Shiffrar, 1999). Furthermore, the learning of unfamiliar motion patterns is better when the joints make articulated movements constrained by an (invisible) skeleton (Jastorff, Kourtzi, & Giese, 2006). These findings suggested the importance of certain relationships between joints in achieving our high sensitivity to biological motion signals. It has, however, so far, been difficult to describe precisely which relationships between joints are most important. Apart from some research indicating that the position of the feet relative to the rest of the body is important (Hirai et al., 2011; Thurman & Lu, 2013b), we know very little about the importance of these cues.

The logistic regression analysis that we have employed in deriving the CIs has the advantage of estimating how multiple joints jointly contribute to determine the judgment of action discrimination. This method therefore identifies relational features that are used by observers and serve as an approximation of relational processing in action perception. This type of analysis goes beyond simple first-order CIs and takes into account higher-order statistics. Our study is the first to look at the correlation between predictors in reconstructing CIs for biological motion (i.e., joints in our experiment), although some previous work has examined covariance (Neri, 2009a), which is related to our correlation analysis, and other types of high-order statistics (Neri, 2004; Neri & Heeger, 2002; Neri, Parker, & Blakemore, 1999) in the context of brightness and orientation perception.

Interestingly, only discrimination of upright actions yielded significant interjoint relation features in the resulting CIs. There were two main types of relational features. The first was the relation associating the contributions of the arms with the hips. We hypothesize that the visual system uses the hip movements to infer the body layout in a 3D world, so that the relative movements of the arms can be calculated in the body-centered coordinates. In addition, the brain identifies actions relying on certain discriminative properties of the relational feature (e.g., wrists below the hips indicate walking, wrists above the hips indicate running). The second type of relational feature (present only in the side view) was the relationship between the two feet. Indeed, the movements of the two feet relative to each other have been reported to be important in identifying the walking direction of an actor (e.g., Casile & Giese, 2005; Thurman & Grossman, 2008), as they produce opponent motion signals. Although it is still speculative to provide functional accounts for relational features, the existence of these features suggests that, indeed, the brain employs relational processing, and a more holistic approach in general, when viewing action in the natural viewing orientation.

Conclusion

Overall, with our technique, we identified critical joint features, foremostly the feet and the wrists, in an action discrimination task of bipedal locomotion. We furthermore showed that many other joints have a significant influence on classification decisions in the side view, suggestive of a form-based template-matching process. Moreover, we found that the visual system employs several relational features, such as the combined information from the two feet, but only in ecologically valid upright body orientations.

Our observations are made with stimuli that are presented for 1 s. This is a duration at which the biological motion analysis seems to rely on motion information predominantly and not on form information (Thurman et al., 2010). It would be of interest to see if shorter trial durations would result in a greater dependence on form information and potentially template-matching mechanisms (Lange & Lappe, 2006). Such shorter trials could potentially manifest a greater reliance on structural features as measured in the side view conditions in our experiments.

Our data were obtained in 1,000 trials, which are greatly fewer trials than in many other CI paradigms in the literature of biological motion research. Obtaining similar results with a hypothesis-driven approach would take a large effort and requires the involvement of many different stimulus manipulations. In addition, the traditional hypothesis-driven approach may also not provide the level of detail that we show in the present study (i.e., comparative contributions of all joints and their relations). Therefore, this new approach is a useful addition to the psychophysical tools available to the researcher interested in biological motion perception.

Our discussion has focused mainly on biological motion tasks that involve walking and running actions, because most research is done on these two actions. However, observers are generally very good at identifying many types human action. Observers can easily recover information about, for example, identity (Cutting & Kozlowski, 1977), emotion (Dittrich et al., 1996; Pollick et al., 2001), and type of actions other than walking (Brown et al., 2005; Dittrich, 1993; Ma et al., 2006; Norman et al., 2004; van Boxtel & Lu, 2011). One further advantage of the hypothesis-free approach that we have employed in the current design is that it is equally well applicable to other types of classification stimulus (e.g., boxing in which the wrists are likely more important; van Boxtel & Lu, 2012) and other tasks (e.g., emotion discrimination).

Keywords: classification images, action perception, action classification, biological motion, double-pass paradigm

Acknowledgments

This project was supported by a grant from the National Science Foundation (NSF BCS-1353391) awarded to H. L. We thank Daniel Lin for help in running the experiments and Steven Thurman for insightful discussions. We also thank two reviewers for their comments.

Commercial relationships: none.

Corresponding author: Jeroen J. A. van Boxtel.

Email: j.j.a.vanboxtel@gmail.com.

Address: Department of Psychology, University of California, Los Angeles, Los Angeles, CA, USA.

Footnote

¹Reverse correlation experiments in psychology are often analyzed with linear classification models and not with logistic regressions. We therefore also calculated classification images based on linear models. Classification images were calculated for each individual, and the statistics were then performed over subjects, which yielded results nearly identical to the logistic regression data in terms of contributions from individual joints. As a further alternative, combining all subjects' data and computing a classification image over the total set of trials, and doing statistics with z-scores obtained from response-shuffled data, also yielded very similar data. However, the classical CI analysis cannot reveal effects of relationships between joints.

References

- Ahumada, A. J., Jr. (2002). Classification image weights and internal noise level estimation. *Journal of Vision*, 2(1):8, 121–131, <http://www.journalofvision.org/content/2/1/8>, doi:10.1167/2.1.8. [PubMed] [Article]
- Ahumada, A. J., & Lovell, J. (1971). Stimulus features in signal detection. *Journal of the Acoustical Society of America*, 49, 1751–1756.
- Barlow, H. B. (1978). The efficiency of detecting changes of density in random dot patterns. *Vision Research*, 18, 637–650.
- Beard, B. L., & Ahumada, A. J., Jr. (1998). A technique to extract relevant image features for visual tasks. In B. E. Rogowitz & T. N. Pappas (Eds.), *Human vision and electronic imaging III, SPIE Proceedings, Vol. 3299* (pp. 79–85).
- Beintema, J. A., & Lappe, M. (2002). Perception of biological motion without local image motion. *Proceedings of the National Academy of Sciences, USA*, 99, 5661–5663.
- Bertenthal, B. I., & Pinto, J. (1994). Global processing of biological motions. *Psychological Science*, 5, 221–224.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, 10, 433–436.
- Brown, W. M., Cronk, L., Grochow, K., Jacobson, A., Liu, C. K., Popovic, Z., & Trivers, R. (2005). Dance reveals symmetry especially in young men. *Nature*, 438(7071), 1148–1150.
- Burgess, A. E., & Colborne, B. (1988). Visual signal detection. IV. Observer inconsistency. *Journal of the Optical Society of America. A, Optics and Image Science*, 5, 617–627.
- Casile, A., & Giese, M. A. (2005). Critical features for the recognition of biological motion. *Journal of Vision*, 5(4):6, 348–360, <http://www.journalofvision.org/content/5/4/6>, doi:10.1167/5.4.6. [PubMed] [Article]
- Chang, D. H., & Troje, N. F. (2009). Acceleration carries the local inversion effect in biological motion perception. *Journal of Vision*, 9(1):19, 1–17, <http://www.journalofvision.org/content/9/1/19>, doi:10.1167/9.1.19. [PubMed] [Article]
- Cutting, J. E., & Kozlowski, L. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, 9, 353–356.
- Diedrich, F. J., & Warren, W. H., Jr. (1995). Why change gaits? Dynamics of the walk-run transition. *Journal of Experimental Psychology: Human Perception and Performance*, 21, 183–202.
- Dittrich, W. H. (1993). Action categories and the perception of biological motion. *Perception*, 22, 15–22.
- 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.
- Dupuis-Roy, N., Fortin, I., Fiset, D., & Gosselin, F. (2009). Uncovering gender discrimination cues in a realistic setting. *Journal of Vision*, 9(2):10, 1–18, <http://www.journalofvision.org/content/9/2/10>, doi:10.1167/9.2.10. [PubMed] [Article]
- Eckstein, M. P., & Ahumada, A. J., Jr. (2002). Classification images: A tool to analyze visual strategies. *Journal of Vision*, 2(1):i, <http://www.journalofvision.org/content/2/1/i>, doi:10.1167/2.1.i. [PubMed] [Article]

- Frontier, S. (1976). Étude de la décroissance des valeurs propres dans une analyse en composantes principales: comparaison avec le modèle du bâton brisé. *Journal of Experimental Marine Biology and Ecology*, 25, 67–75.
- Giese, M. A., & Lappe, M. (2002). Measurement of generalization fields for the recognition of biological motion. *Vision Research*, 42, 1847–1858.
- Giese, M. A., & Poggio, T. (2000). Morphable models for the analysis and synthesis of complex motion patterns. *International Journal of Computer Vision*, 38, 59–73.
- Gold, J. M., Murray, R. F., Bennett, P. J., & Sekuler, A. B. (2000). Deriving behavioural receptive fields for visually completed contours. *Current Biology*, 10, 663–666.
- Gold, J. M., Tadin, D., Cook, S. C., & Blake, R. (2008). The efficiency of biological motion perception. *Perception & Psychophysics*, 70, 88–95.
- Heeger, D. J., Boynton, G. M., Demb, J. B., Seidemann, E., & Newsome, W. T. (1999). Motion opponency in visual cortex. *Journal of Neuroscience*, 19, 7162–7174.
- Hill, H., & Pollick, F. E. (2000). Exaggerating temporal differences enhances recognition of individuals from point light displays. *Psychological Science*, 11, 223–228.
- Hirai, M., Chang, D. H., Saunders, D. R., & Troje, N. F. (2011). Body configuration modulates the usage of local cues to direction in biological-motion perception. *Psychological Science*, 22, 1543–1549.
- Hiris, E. (2007). Detection of biological and nonbiological motion. *Journal of Vision*, 7(12):4, 1–16, <http://www.journalofvision.org/content/7/12/4>, doi:10.1167/7.12.4. [PubMed] [Article]
- Jastorff, J., Kourtzi, Z., & Giese, M. A. (2006). Learning to discriminate complex movements: Biological versus artificial trajectories. *Journal of Vision*, 6(8):3, 791–804, <http://www.journalofvision.org/content/6/8/3>, doi:10.1167/6.8.3. [PubMed] [Article]
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14, 201–211.
- Kontsevich, L. L., & Tyler, C. W. (2004). What makes Mona Lisa smile? *Vision Research*, 44, 1493–1498.
- Lange, J., & Lappe, M. (2006). A model of biological motion perception from configural form cues. *Journal of Neuroscience*, 26, 2894–2906.
- Li, R., Levi, D., & Klein, S. (2004). Perceptual learning improves efficiency by re-tuning the decision ‘template’ for position discrimination. *Nature Neuroscience*, 7, 178–183.
- Lu, H., & Liu, Z. (2006). Computing dynamic classification images from correlation maps. *Journal of Vision*, 6(4):12, 475–483, <http://www.journalofvision.org/content/6/4/12>, doi:10.1167/6.4.12. [PubMed] [Article]
- Lu, Z. L., & Doshier, B. A. (2008). Characterizing observers using external noise and observer models: Assessing internal representations with external noise. *Psychological Review*, 115, 44–82.
- Ma, Y., Paterson, H. M., & Pollick, F. E. (2006). A motion capture library for the study of identity, gender, and emotion perception from biological motion. *Behavior Research Methods*, 38, 134–141.
- Mark Williams, A., Huys, R., Canal-Bruland, R., & Hagemann, N. (2009). The dynamical information underpinning anticipation skill. *Human Movement Science*, 28, 362–370.
- Mather, G., Radford, K., & West, S. (1992). Low-level visual processing of biological motion. *Proceedings Biological Science*, 249, 149–155.
- Neri, P. (2004). Estimation of nonlinear psychophysical kernels. *Journal of Vision*, 4(2):2, 82–91, <http://www.journalofvision.org/content/4/2/2>, doi:10.1167/4.2.2. [PubMed] [Article]
- Neri, P. (2009a). Nonlinear characterization of a simple process in human vision. *Journal of Vision*, 9(12):1, 1–29, <http://www.journalofvision.org/content/9/12/1>, doi:10.1167/9.12.1. [PubMed] [Article]
- Neri, P. (2009b). Wholes and subparts in visual processing of human agency. *Proceedings Biological Sciences*, 276, 861–869.
- Neri, P. (2010). How inherently noisy is human sensory processing? *Psychonomic Bulletin & Review*, 17, 802–808.
- Neri, P., & Heeger, D. J. (2002). Spatiotemporal mechanisms for detecting and identifying image features in human vision. *Nature Neuroscience*, 5, 812–816.
- Neri, P., & Levi, D. M. (2006). Receptive versus perceptive fields from the reverse-correlation viewpoint. *Vision Research*, 46, 2465–2474.
- Neri, P., Parker, A. J., & Blakemore, C. (1999). Probing the human stereoscopic system with reverse correlation. *Nature*, 401, 695–698.
- Norman, J. F., Payton, S. M., Long, J. R., & Hawkes, L. M. (2004). Aging and the perception of biological motion. *Psychology and Aging*, 19, 219–225.
- Pavlova, M., & Sokolov, A. (2000). Orientation

- specificity in biological motion perception. *Perception & Psychophysics*, *62*, 889–899.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, *10*, 437–442.
- Perrett, D. I., Smith, P. A., Mistlin, A. J., Chitty, A. J., Head, A. S., Potter, D. D., . . . Jeeves, M. A. (1985). Visual analysis of body movements by neurones in the temporal cortex of the macaque monkey: A preliminary report. *Behavioral Brain Research*, *16*, 153–170.
- Pinto, J., & Shiffrar, M. (1999). Subconfigurations of the human form in the perception of biological motion displays. *Acta Psychologica*, *102*, 293–318.
- Pollick, F. E., Lestou, V., Ryu, J., & Cho, S. B. (2002). Estimating the efficiency of recognizing gender and affect from biological motion. *Vision Research*, *42*, 2345–2355.
- Pollick, F. E., Paterson, H. M., Bruderlin, A., & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition*, *82*, B51–B61.
- Proffitt, D. R., & Bertenthal, B. I. (1990). Converging operations revisited: Assessing what infants perceive using discrimination measures. *Perception & Psychophysics*, *47*, 1–11.
- Ringach, D. L., Sapiro, G., & Shapley, R. (1997). A subspace reverse-correlation technique for the study of visual neurons. *Vision Research*, *37*, 2455–2464.
- Roether, C. L., Omlor, L., Christensen, A., & Giese, M. A. (2009). Critical features for the perception of emotion from gait. *Journal of Vision*, *9*(6):15, 1–32, <http://www.journalofvision.org/content/9/6/15>, doi:10.1167/9.6.15. [PubMed] [Article]
- Saunders, D. R., Suchan, J., & Troje, N. F. (2009). Off on the wrong foot: Local features in biological motion. *Perception*, *38*, 522–532.
- Sumi, S. (1984). Upside-down presentation of the Johansson moving light-spot pattern. *Perception*, *13*, 283–286.
- Thornton, I. A., Pinto, J., & Shiffrar, M. (1998). The visual perception of human locomotion. *Cognitive Neuropsychology*, *15*, 535–552.
- Thurman, S. M., Giese, M. A., & Grossman, E. D. (2010). Perceptual and computational analysis of critical features for biological motion. *Journal of Vision*, *10*(12):15, 1–14, <http://www.journalofvision.org/content/10/12/15>, doi:10.1167/10.12.15. [PubMed] [Article]
- Thurman, S. M., & Grossman, E. D. (2008). Temporal “bubbles” reveal key features for point-light biological motion perception. *Journal of Vision*, *8*(3):28, 1–11, <http://www.journalofvision.org/content/8/3/28>, doi:10.1167/8.3.28. [PubMed] [Article]
- Thurman, S. M., & Lu, H. (2013a). Complex interactions between spatial, orientation, and motion cues for biological motion perception across visual space. *Journal of Vision*, *13*(2):8, 1–18, <http://www.journalofvision.org/content/13/2/8>, doi:10.1167/13.2.8. [PubMed] [Article]
- Thurman, S. M., & Lu, H. (2013b). Physical and biological constraints govern perceived animacy of scrambled human forms. *Psychological Science*, *24*, 1133–1141.
- Tjan, B. S., Braje, W. L., Legge, G. E., & Kersten, D. (1995). Human efficiency for recognizing 3-D objects in luminance noise. *Vision Research*, *35*, 3053–3069.
- Troje, N. F., & Westhoff, C. (2006). The inversion effect in biological motion perception: Evidence for a “life detector”? *Current Biology*, *16*, 821–824.
- van Boxtel, J. J., & Lu, H. (2011). Visual search by action category. *Journal of Vision*, *11*(7):19, 1–14, <http://www.journalofvision.org/content/11/7/19>, doi:10.1167/11.7.19. [PubMed] [Article]
- van Boxtel, J. J., & Lu, H. (2012). Signature movements lead to efficient search for threatening actions. *PLoS ONE*, *7*(5), e37085.
- van Boxtel, J. J., & Lu, H. (2013). A biological motion toolbox for reading, displaying, and manipulating motion capture data in research settings. *Journal of Vision*, *13*(12):7, 1–16, <http://www.journalofvision.org/content/13/12/7>, doi:10.1167/13.12.7. [PubMed] [Article]
- van Boxtel, J. J. A., & Lu, H. (2013). Impaired global, and compensatory local, biological motion processing in people with high levels of autistic traits. *Frontiers in Psychology*, *4*, 209.
- Vangeneugden, J., De Maziere, P. A., Van Hulle, M. M., Jaeggli, T., Van Gool, L., & Vogels, R. (2011). Distinct mechanisms for coding of visual actions in macaque temporal cortex. *Journal of Neuroscience*, *31*, 385–401.
- Vangeneugden, J., Pollick, F., & Vogels, R. (2009). Functional differentiation of macaque visual temporal cortical neurons using a parametric action space. *Cerebral Cortex*, *19*, 593–611.
- Victor, J. D. (2005). Analyzing receptive fields, classification images and functional images: Challenges with opportunities for synergy. *Nature Neuroscience*, *8*, 1651–1656.