In: Shipley, T. F. and Zacks (eds.) Understanding Events: From Perception to Action. Oxford University Press, 2008.

Erratum: Equations (4) and (5) on page 317 are not right. The correct versions are:

$$w_i = v_0 + \sum_{j=1}^{m} k_{i,j} v_j \quad (4)$$

$$W = V_0 + VK \quad (5)$$

12

Retrieving Information from Human Movement Patterns

NIKOLAUS F. TROJE

Biological motion—that is, the movement patterns of animals and humans—provides a rich source of information that helps us to quickly and reliably detect the presence of another living being; to identify it as a predator, prey, or conspecific; and to infer its actions and intentions in order to respond with adequate behavior. Once we know that we are being confronted with another person, we are able to use motion as a source of information about identity, gender, age, emotional state, and personality traits and as a complex means for signaling and communications.

More than 30 years ago, the Swedish psychologist Gunnar Johansson (1973) introduced a stimulus into experimental psychology that allows us to disentangle to a large degree the information contained in the kinematics of a moving body from other sources of information about action and identity. His work showed that a few light dots placed strategically on a moving human or animal body are instantaneously organized into the coherent percept of a living creature (also see Chapter 11 in this volume). The observation goes back to the earlier work of the pioneers of cinematography (Muybridge, 1887/1979) and biomechanics (Marey, 1895/1972), but it was Johansson (1973) who first

appreciated the significance of the tremendous saliency of biological motion point-light displays, and the effortless perceptual organization with which the visual system responds to them. Fewer than 10 isolated dots and display times of 200 ms are sufficient for a vivid percept of the articulated structure of a human body (Johansson, 1976).

Subsequent research concentrated on several aspects of this general phenomenon. It was shown that biological motion perception is very robust in the presence of many types of distracting masks (Cutting, Moore, & Morrison, 1988) and that it reveals more than just the presence of a person: point-light displays convey information about sex (Barclay, Cutting, & Kozlowski, 1978; Cutting, Proffitt, & Kozlowski, 1978; Kozlowski & Cutting, 1978) and identity of an agent (Cutting & Kozlowski, 1977), as well as emotional attributes (Dittrich, Troscianko, Lea, & Morgan, 1996; Pollick, Paterson, Bruderlin, & Sanford, 2001). Infants can perceive biological motion (Bertenthal, Proffitt, & Kramer, 1987; Fox & McDaniel, 1982), and it has been shown that at least pigeons and cats respond specifically to point-light displays (Blake, 1993; Dittrich, Lea, Barrett, & Gurr, 1998).

While the early work mainly concentrated on demonstrating the abilities of the visual system in processing biological motion, more recent studies have helped us understand how this is being achieved and which parts of the brain are involved. Recording from single cells in macaque cortex, Oram and Perrett (1994) first identified structures in the upper bank of the superior temporal sulcus (STS) as selectively responsive to human form and motion. A number of more recent brain imaging studies corroborate this finding and show that the posterior part of STS (STSp) is particularly active when looking at point-light displays of an upright human walker (Bonda, Petrides, Ostry, & Evans, 1996; Grossman, Blake, & Kim, 2004; Grossman et al., 2000; Peuskens, Vanrie, Verfaillie, & Orban, 2005). While STS is clearly responsive to biological motion, it is not clear how specific this area is. Stimuli such as speech (Beauchamp, 2005) or the sound of footsteps (Bidet-Caulet, Voisin, Bertrand, & Fonlupt, 2005), as well as motion confined to specific limbs, the eyes, or the mouth (Grezes, Costes, & Decety, 1998; Puce, Allison, Bentin, Gore, & McCarthy, 1998), also result in STSp activation. Besides STSp, other areas have been identified that are responsive to biological motion. They include the ventral surface of the temporal lobe (Vaina, Solomon, Chowdhury, Sinha, & Belliveau, 2001),

the fusiform gyrus (Beauchamp, Lee, Haxby, & Martin, 2002), and the fusiform face area (Grossman & Blake, 2002; Peelen & Downing, 2005). For all these areas, it is rather unclear if they respond specifically to human motion or if they are triggered generally by biological motion. Very few imaging studies have contrasted responses to representations of humans versus nonhumans, and none of these used standard biological motion point-light displays (Buccino et al., 2004; Downing, Jiang, Shuman, & Kanwisher, 2001).

For a long time, biological motion has been treated as a single phenomenon. Only during the past few years has it become obvious that there are a number of different mechanisms involved that need to be distinguished both conceptually as well as experimentally. Here I am suggesting at least four different stages of information processing involved with biological motion perception.

1. **Detection of animate motion.** A fast and reliable system is required to detect the presence of an animal in the visual environment. Ideally, this mechanism should be independent of the particular nature of the animal, and in particular independent of its shape. It should respond to biological motion in the whole visual field, including the visual periphery. The evolutionary significance of such an early "life detector" is obvious. It is required either to trigger fast behavioral responses (flight, attack) or to guide attention to potentially threatening or otherwise interesting events. Troje and Westhoff (2006) identified the ballistic movements of the limbs of a terrestrial animal to provide such an invariant. The cue seems to work well not only for foveal vision but also in the visual periphery, and probably directs attention to an event of potentially vital significance. The underlying visual filter mechanism shows a pronounced inversion effect: if presented upside-down, our visual system is no longer able to retrieve information from the local motion of the limbs. The visual filter is expected to be evolutionarily old, innate rather than learned, and shared by other animals. Behavioral experiments on visually naïve, newly hatched chicks suggest that they in fact use the same cue to identify the object of filial imprinting (Vallortigara & Regolin, 2006; Vallortigara, Regolin, & Marconato, 2005).

2. **Structure from motion.** Once a living creature is detected, its movements can be used to perceptually organize it into a coherent, articulated body structure, resulting in "basic-level" (Rosch, 1988) agent recognition (e.g., Is this a human, a cat, a bird?). This mechanism does not work very well in the visual periphery (Ikeda, Blake, & Watanabe, 2005) and probably requires attention (Cavanagh, Labianca, & Thornton, 2001; Thornton, Rensink, & Shiffrar, 2002). In contrast to the early "life detection" stage, it requires learning and individual experience (Jastorff, Kourtzi, & Giese, 2006). It is also subject to an inversion effect, which, however, is independent of the one operating on the "life detection" mechanism and is instead similar to the orientation dependency of configural processing observed in face recognition (Farah, Tanaka, & Drain, 1995).
3. **Action perception.** On this level, structural and kinematic information is integrated into a system that classifies and categorizes actions and events. Ideally, efficient classification on this level should be invariant to actor, viewpoint, and the particular style of the action. Many of the chapters in this section of the book specifically address this processing level (see especially Chapters 10 and 11).
4. **Style recognition.** Once both agent and action are identified, pattern recognition at a "subordinate" (Rosch, 1988) level helps to retrieve further information about the details of both. For instance, once we know we are confronted with a human walker (rather than, say, a hunting tiger), we are able to use motion as a source of information about individual identity, gender, age, emotional state, and personality traits, and as a complex means for signaling and communications. Depending on the particular property, the results of initial data processing required to characterize and isolate diagnostic features might eventually feed into different neuronal circuits, and in that respect "style recognition" might not be due to a single mechanism but to several. Yet, at least from a computational point of view, it is likely that all of them share certain processing principles.

For the remainder of this chapter, I will concentrate on style recognition. The ability of the human visual system to sensitively detect

and adequately interpret subtle nuances in the way people move is both a prerequisite and a consequence of the fact that the complex social structures characterizing our species require us to identify one another individually and to attribute emotion, personality, and intentions to our peers.

A number of different approaches to understanding style in human movement come both from the computer-vision community and from experimental psychology, and a large part of this work is well summarized in Chapter 11. In the present chapter, I want to present and discuss a particular computational framework that we used to retrieve stylistic information from visual human locomotion patterns over the past years. It was first developed to identify and analyze sex-specific differences between walkers (Troje, 2002a). We then changed and further improved the algorithm and applied it to a number of different problems and questions in the context of pattern recognition from biological motion. In the next section, I will first outline the general framework and then provide the details of the algorithm. In the third section, I will summarize some of the studies in which we applied the algorithm. In the final section, I will discuss the role of the proposed framework in understanding the very complex class of stimuli that our visual system copes with so easily, its value as a model for human perception, and potential ways to generalize and improve it.

A Framework for the Analysis and Synthesis of Human Walking Patterns

Our approach to understanding the mechanisms underlying biological motion perception focuses on the information provided by the stimulus itself. Understanding how information is encoded in biological motion patterns is a prerequisite for designing artificial vision systems, but is also helpful for understanding biological vision systems by means of "reverse engineering." How can we possibly retrieve structure from the complex spatiotemporal patterns of animate motion? Approaching biological motion perception as a pattern-recognition problem can teach us about principles and constraints that any system, regardless of whether it is artificial or biological, has to cope with when analyzing biological motion patterns.

The general idea of our approach is as follows. Starting with data obtained by means of a motion capture system (i.e., with the 3-D trajectories of discrete points on a person's body), the first step of the subsequent data processing is to transform the data into a representation that would allow us to apply standard methods from linear statistics and pattern recognition. Such representations have been termed "morphable models" (Giese & Poggio, 2000; Jones & Poggio, 1999; Shelton, 2000) in the computer-vision community, expressing the fact that the linear transition from one item to a second item of the data set represents a well-defined, smooth metamorphosis such that all intermediate stages maintain the structural characteristics defining the object class and are therefore qualitatively indistinguishable from the start and end points. Other terms that have been used in object recognition for similar kinds of models are "linear object classes" (Vetter & Poggio, 1997) and, in the context of human face recognition, "correspondence-based representations" (Troje & Vetter, 1998; Vetter & Troje, 1997). This latter term focuses on the fact that morphable models rely on establishing correspondence between features across the data set, resulting in a separation of the overall information into range-specific information on the one hand and domain-specific information on the other (Ramsay & Silverman, 1997). We also use the term "linearization" for the nonlinear transformation that is required to establish a representation that then enables us to treat the data as objects in linear space.

Linearization of the walker data mainly involves matching them in terms of frequency and phase. We do this by first computing the Fourier transform and then matching the data directly in the frequency domain. Note that this is slightly different from the way we have described data processing in earlier work (Troje, 2002a), where we used principal components analysis (PCA) to reduce dimensionality of the set of poses of a single walker. We found that the weights of the resulting Eigenposes (that is, the characteristic poses spanning the space of poses of an individual walker) vary sinusoidally with time. The decomposition resulting from PCA is therefore very similar to the one resulting from Fourier analysis. The main difference is that the roles of the terms that represent the basis functions and the terms that constitute the coefficients on these basis functions are interchanged (Troje, 2002b).

Once the data are linearized we apply PCA—however, this time we apply it not to the pose space but to the whole Fourier-based representations of a set of walkers. This reduces dimensionality of the linear walker space to the degree that the number of dimensions is much smaller than the number of data points that we use to establish this space. This step is important to avoid overfitting in the subsequent classification and to eventually construct a classifier that has predictive value.

The classifier itself is a very simple one. Based on the low-dimensional, linear space resulting from the previous two steps, we compute a linear discriminant function (LDF) by means of linear regression of the class indicator (e.g., indicating whether a particular walker in a training set is male or female) on the projections of the walkers into the resulting space.

One interesting feature of our approach is the fact that the transformations that map the time series of original motion capture data onto the morphable, low-dimensional walker space are more or less lossless and are therefore invertible. Consequently, any point in the morphable space—even points that do not correspond to original walkers—can be transformed back into a time series of marker positions and visualized as a point-light display. We will use this to exaggerate and caricature the set of diagnostic features that the classifier extracts, and to generate walking patterns with the respective properties and attributes.

The procedures described in the present study contain elements of earlier work on parameterizations of animate motion patterns (Bruderlin & Williams, 1995; Giese & Poggio, 2000; Guo & Roberge, 1996; Rose, Bodenheimer, & Cohen, 1998; Unuma, Anjyo, & Takeuchi, 1995; Urtasun, Glardon, Ronan, Thalmann, & Fua, 2004; Witkin & Popovic, 1995). Perhaps the most important one in the context of this paper is Unuma et al.'s (1995) study, which showed that blending between different human movements works much better in the frequency domain. At least for periodic motions, such as most locomotion patterns, Fourier decomposition can be used to achieve efficient, low-dimensional, linear decompositions. The fact that PCA applied to a time series of poses of a single walking person basically results in a discrete Fourier decomposition demonstrates that Fourier decomposition of walking data is nearly optimal in terms of explaining a maximum of variance with a minimum number of components (Troje, 2002b).

Data Collection

Most of our current work is based on data obtained by means of optical motion capture (Vicon 512, Oxford Metrics) from human subjects walking either on a treadmill or over ground. An array of 9 to 12 high-speed (120 Hz) cameras tracks the 3-D position of small (14 mm in diameter), passively reflecting markers with a spatial acuity in the order of 1 mm. Typically, participants wear swimsuits, and most of the markers are taped directly to their skin. Others, like the markers for the head, the ankles, and the wrists, are fixed to elastic bands, and the ones on the feet are taped onto the subjects' shoes. Currently, we are using a set of 41 markers (often referred to as the Helen-Hayes marker set; Davis, Ounpuu, Tyburski, & Gage, 1991). This set is designed such that, together with a few anthropometric measurements (such as the width of the knee, elbow, ankle, etc.), it provides the input to a biomechanical model that outputs accurate estimates for the location of the major joints of the human body. For most of our work, we use a set of 15 derived, "virtual" markers. They are located at shoulder joints, elbows, wrists, hip joints, knees, and ankles, and at the centers of the pelvis, clavicles, and head. The motion data that provide the input for the subsequent processing are therefore time series of poses (sampled at 120 Hz), each consisting of the 3-D Cartesian coordinates of 15 virtual markers. We call the 45-dimensional vector specifying their current location in space a "pose."

Whether data were collected from subjects walking on a treadmill or freely on the ground, we always asked our participants to walk for at least 5 minutes before we started data collection. On the treadmill, we allowed them to try different belt speeds to make sure that they felt as comfortable as possible. We did not tell the participants when actual data collection started. On the treadmill, we recorded a sequence of at least 10 full gait cycles. The volume available for free over-ground walking covered about four full gait cycles, and we typically recorded four passes throughout this volume for every participant.

Data Processing

The walk of an individual subject can be regarded as a time series of poses. Each pose can be described in terms of the position of the

15 markers in 3-D space. A single pose is therefore represented by a 45-dimensional vector:

$$P = (m1_x, m1_y, m1_z, m2_x, \ldots m15_z)^T \tag{1}$$

The time series of poses of a particular walker i can be decomposed into a discrete second-order Fourier series,

$$\begin{aligned} P_i(t) = {} & p_{i,0} + p_{i,1}\sin(\omega_i t) + q_{i,1}\cos(\omega_i t) + p_{i,2}\sin(2\omega_i t) \\ & + q_{i,2}\cos(2\omega_i t) + \mathrm{err}_i \end{aligned} \tag{2}$$

where ω is the angular frequency derived from the gait frequency f as $\omega = 2\Pi f$. The first term $p_{i,0}$ describes the time-invariant average pose of walker i. It contains anthropometric structural information, for instance about the length of the limbs, the width of the shoulders, etc. The next two terms specify amplitudes and phases of the fundamental frequency, and the final two terms contain information about the second harmonic. Pairs of $p_{i,j}$ and $q_{i,j}$ can be translated directly into amplitudes $a_{i,j}$ and phases $\varphi_{i,j}$.

$$a_{i,j} = \sqrt{p_{i,j}^2 + q_{i,j}^2}, \quad \varphi_{i,j} = \arctan\left(\frac{q_{i,j}}{p_{i,j}}\right), \tag{3}$$

Here, Fourier analysis serves two purposes. First, it very effectively reduces redundancy and therefore provides compression of the data. For walking, the power carried by the residual term err in Equation 2 is less than 1% of the overall variance of the input data, and we usually discard it in all further computations.

Second, we use the Fourier representation to register the data in order to define correspondence between individual walking sequences. This is done by simply adjusting the phase and the frequency of the individual walking sequences. While the frequency of the walk is expressed in terms of the fundamental frequency ω_i, the absolute phase of the sequence and the relative phases between the 15 markers are contained in the relative contributions to the sine and the cosine terms (see Eq. 3). The absolute phase of a walking sequence depends only on the time at

which we started data collection and contains no further information. We therefore adjusted the absolute phase of all sequences such that the average phase angle of the two ankle markers is 0 degrees.

A walk of a particular subject i is now approximated by specifying the average pose $p_{i,0}$, the four characteristic poses $p_{i,1}$, $q_{i,1}$, $p_{i,2}$, and $q_{i,2}$, and the fundamental frequency ω_i. The average pose and characteristic poses are all 45-dimensional vectors, while the fundamental frequency is a scalar. Thus, the dimensionality of the model at this stage is $5 \times 45 + 1 = 226$.

Although this number already reflects a considerable reduction in dimensionality as compared to the raw motion capture data, we expect the number of effective degrees of freedom within the database to be much smaller. For classification purposes it is necessary to reduce the dimensionality of the representation such that the number of dimensions becomes much smaller than the number of items represented in the resulting space.

The advantage of the above "linearized" representation (Eq. 2) is that it makes it possible to successfully apply linear operations to the set of motion data. At this stage, our representation has become a morphable model. Linear combinations of existing walking patterns result in new walking patterns that meaningfully represent the transitions between the constituting patterns (Troje, 2002a; Unuma et al., 1995). We can treat the 226-dimensional vector describing the walk wi of walker i as an object in linear space.

Classification

This representation also makes it possible to use PCA in order to further reduce dimensionality. Applying PCA to the set of walkers W results in a decomposition of each walker into an average walker v_0 and a weighted sum of Eigenwalkers v_j,

$$w_i = \sum_{j=1}^{m} k_{i,j} v_j \tag{4}$$

or, in matrix notation,

$$W = V_0 = VK \tag{5}$$

V_0 denotes a matrix with the average walker v_0 in each column. The matrix V contains the m Eigenwalkers as column vectors v_j. Matrix K contains the weights (or the coefficients) ki,j and is obtained by solving the linear equation system

$$VK = W - V_0 \tag{6}$$

Given a set of n walkers, this procedure yields a total of $n-1$ Eigenwalkers. However, the variance covered by the first m Eigenwalkers is generally much larger than the fraction m/n of the overall variance. For instance, with the set of 100 walkers that we currently use, only four components are required to cover 50% of the overall variance, and 22 are required to cover 90%. The final choice of the number of Eigenwalkers m depends on the particular application. For classification purposes, it is recommended to use a relatively low number, resulting in better generalization. With a set of 100 walkers, we typically use 10 principal components for classification. For visualization purposes a larger number might be more suitable, since reconstruction is more accurate.[1]

Within the space spanned by the first m Eigenwalkers, a linear discriminant function can now be computed by simply regressing the class indicator (or any other variable quantifying an attribute of interest) on the projections of the walkers in the Eigenwalker space. This is achieved by finding the best (according to a least-square criterion) solution d of the overdetermined linear system

$$K^T d = r \tag{7}$$

K^T is the transpose of matrix K, which contains the coefficients of each walker in the Eigenwalker space (Eq. 5). r is the column vector containing the n values r_i indicating the class to which walker i belongs (e.g., r_i

1. In our implementations, rather than submitting the matrix $W - V_0$ to the PCA, we first normalized each row of this matrix to get unit variance for each of them. The reasons for doing this, and the consequences it has for the subsequent computations and the way the resulting principal components are used, are beyond the scope of this chapter, and I chose to omit this detail here.

equals 1 if walker *i* is male and –1 if the walker is female) or another attribute that encodes the property of interest (e.g., a Likert scale rating given by an observer). The resulting column vector *d* then contains the coefficients of the linear discriminant function in the Eigenwalker space best accounting for the gender of the walkers.

The amount of variance explained by the regression (R^2) depends on the number of Eigenwalkers that have been used to span the walker space (*m* in Eq. 4). If all Eigenwalkers are used ($m = n - 1$), the value of R^2 will equal 1. However, in this particular case the predictive value of the classifier will be very low. The R^2 value, therefore, is not a good statistic to assess the quality of the classifier. A more useful way to do this is based on a leave-one-out cross-validation. A single walker is taken out, the discriminant function is computed on the set of the other $n - 1$ data samples, and the remaining walker is then used to test the classifier. This procedure is repeated for every single data sample. The classifier can then be evaluated in terms of the percentage of misclassifications.

Feature Extraction

The vector *d* in Equation 7 describes the discriminant function in the low-dimensional space spanned by the first *m* Eigenwalkers. Given the matrix *V* containing the Eigenwalkers themselves, the corresponding discriminant function in the 226-dimensional representation is revealed as

$$v_d = Vd \tag{8}$$

The discriminant walker v_d has the same format as any of the input walkers ωi (see Eq. 4). It can therefore be decomposed into its components according to Equation 2. Note that all of the components describe increments—that is, positive or negative additive terms that modify the average walker v_0. As described above, $p_{d,0}$ encodes structural, anthropomorphic features, whereas $p_{d,j}$ and $q_{d,j}$ represent kinematic information. Particularly, if the individual Fourier terms are transformed into amplitudes and phases (Eq. 3), the contribution of single body parts can also be quantified. Markers that are associated with large numbers in the discriminant walker vd are strongly affected by the attribute of

interest, whereas those associated with small numbers do not change as the attribute changes.

Visualization

The decomposition of the time series of pose data into its Fourier components (Eq. 2) is invertible. A walking pattern that has been manipulated in the frequency domain can be transformed back into a time series of poses, which in turn can be animated in terms of a point-light display. This is a very useful property that helps to create well-defined stimuli for psychophysical experiments and provides tools to explore the nature of the discriminant function and linear classifiers described above. For instance, walkers w_α corresponding to a point that is α standard deviations away from the mean walker are represented as

$$w_\alpha = v_0 + \alpha v_d \tag{9}$$

As above, v_d denotes the discriminant walker and v_0 is the average walker. As α changes from negative to positive values, an animation of the walker v_α appears to change with regard to the attribute on which the classification was based. Large positive or negative α values can be used to generate exaggerated caricatures that help visualize these attributes.

An interesting additional feature of the methodology described here is the option to apply it to walking data that are reduced in information by normalizing it with respect to certain properties while retaining diagnostic information only in others. Below, I will describe a study in which we did that to explore the role of static versus kinematic information for sex classification. Similar manipulations have been used to investigate which parts of the overall information are being used for person identification (Troje, Westhoff, & Lavrov, 2005).

Examples

Sex Classification

Men and women show different walking patterns, and the human visual system is well able to distinguish between them (Barclay et al., 1978;

Cutting et al., 1978; Kozlowski & Cutting, 1977; Mather & Murdoch, 1994). A number of different features have been suggested that are possible candidates for conveying sex-specific diagnostic information. One of them was the "center of moment" (CoM) of the upper body. In an attempt to identify the point of maximal torsion, Cutting (1978) approximated the CoM as the point at which the diagonal lines connecting the shoulder with the contralateral hip intersect. This point is higher in women than in men. Even though the term "center of moment" suggests that this is a dynamic feature, its definition is basically structural. Cutting (1978) demonstrated that the CoM indeed affects sex classification, but later it was shown that this is the case only if no other information is available. A second cue that was investigated by Mather and Murdoch (1994) is the lateral sway of the upper body, which is more pronounced in male walkers than in female walkers. In experiments in which both cues were set into conflict, lateral body sway entirely dominated the CoM.

We applied the framework outlined above to the sex-classification problem (Troje, 2002a). The analysis was based on 20 male and 20 female walkers recorded while walking on a treadmill. The best classifier was based on the first four Eigenwalkers and produced only three misclassifications (out of 40 items), corresponding to an error rate of 7.5%. We then conducted the same analysis using only parts of the overall information. Rather than presenting all the available information to the classifier, we first normalized the data with respect to the size of the walkers and then ran classifications based either only on the remaining structural information ($p_{i,0}$ in Eq. 2) or only on the kinematic information ($p_{i,1}$, $q_{i,1}$, $p_{i,2}$, $q_{i,2}$). It turned out that size alone was a very good predictor for the sex of the walker. When normalized for size, the classification error increased to 17%. Depriving the data of other structural information did not have any effect; however, when kinematic information was removed, misclassifications further increased to 27% (Fig. 12.1). The finding that kinematic information is much more informative than structural information confirms Mather and Murdoch's (1994) results and suggests that the CoM plays only a minor role in sex classification from point-light displays.

To what extent does this result predict how the human visual system processes gender information from walking? We tested this by creating walker stimuli that were normalized with respect to either their structure (kinematics-only) or their kinematics (structure-only) (Troje, 2002a). For

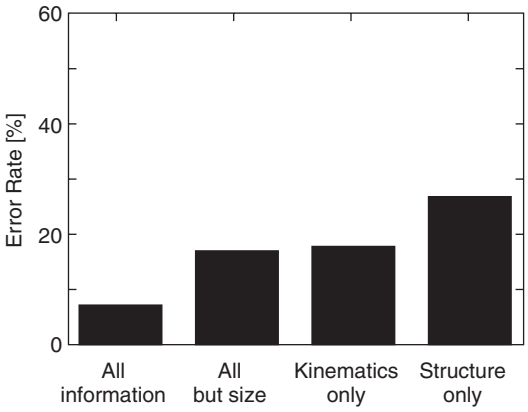

FIGURE 12.1. Percentage errors produced by the linear sex classifier when provided with all available information, size-normalized data, kinematic information only, or structural information only. Note that the information available to the classifier is three-dimensional.

the kinematics-only stimuli, for each individual walker i we replaced the component $p_{i,0}$ (Eq. 2) with the average of these components computed over all 40 walkers. Similarly, for the structure-only stimuli, we replaced the walker-specific components $p_{i,1}$, $q_{i,1}$, $p_{i,2}$, $q_{i,2}$ with their population averages. We then displayed these walkers as point-light displays from three different viewpoints and asked observers to guess the sex of the walkers. The results confirmed that kinematic information is more important than static, structural information for this task (Fig. 12.2). The effect of these manipulations is relatively small when the walkers are shown in frontal view but becomes very substantial for the half-profile and profile views. The results also confirm earlier findings (Mather & Murdoch, 1994) that sex-classification performance is better when walkers are shown in frontal view as compared to the half-profile and profile views.

The differences between male and female walkers can be visualized by animating point-light displays according to Equation 9. Animations of a sex-discriminant function based on the 40 walkers discussed here can be viewed at http://biomotionlab.ca/Demos/BMLgender.html. Visual inspection reveals a number of features that change between men and women. Several differences in the fronto-parallel plane are due to pose and anthropometric structure. Men have wider shoulders and slimmer hips, and their elbows are held farther away from their bodies. In terms of the kinematics, we observe that the upper body of the male shows

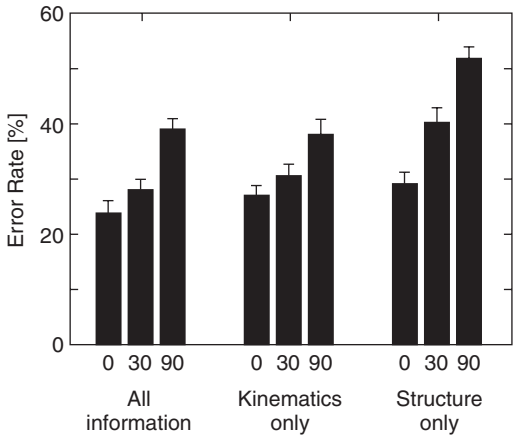

FIGURE 12.2. Percentage error rates (means and standard errors) of human observers who tried to determine the sex of point-light walkers. Point-light displays were shown from three different viewpoints (0 degrees: frontal view; 30 degrees: half-profile view; 90 degrees: profile view). Different groups of observers saw walkers that contained all available information (except size, which was normalized in all displays), kinematic information only, or structural information only. Data are based on 24 observers.

more pronounced lateral sway and that the movement of the hips, though not much larger in amplitude, is different in terms of the phase of its rotation with respect to the other parts of the body. Particularly in the exaggerated animations, it can be seen that the vertical movement of the hip is in counterphase with the vertical motion of knee and foot of the ipsilateral leg in women, while the hip moves almost in phase with the ipsilateral leg in men.

In contrast to previous work, our approach is not primarily hypothesis driven. Rather than focusing on a particular cue (e.g., CoM or lateral body sway), the discriminant function picks up on any feature that distinguishes between male and female walkers, and particularly on the correlations between them. Yet particular hypotheses can be tested and quantified by inspecting the numbers contained in the average walker v_0 and the discriminant function v_d (Eq. 9). For instance, in the average walker, the distance between the two shoulder markers is 350 mm and the distance between the two hip markers is 190 mm. The corresponding differences in the discriminant function are 19 mm for the shoulder and –5 mm for the hip, which means that a walker at a distance of 1 standard deviation (std) into the male part of the space has shoulders that are 19 mm wider and

hips that are 5 mm narrower than in the average walker. Compared to a walker representing a point at a distance of 1 std into the female part of the space, the male walker's shoulders are 38 mm wider and his hips are 10 mm narrower. Transforming the sine and cosine terms of Equation 2 into amplitudes and phases (Eq. 3), we can determine the difference in amplitude of the lateral movement of the shoulders. While the shoulders sway with an amplitude of 20 mm on average (that is, a 40 mm difference between the leftmost and rightmost position) across all 40 walkers, a 1 std male walker sways with an amplitude that is 2.5 mm larger, and a 1 std female walker sways with an amplitude that is 2.5 mm smaller.

Retrieving Information about Other Attributes

In our previous example, the entry r_i of vector r in Equation 7 indicated the sex of walker i. As sex is a binary property, we used the number +1 (for male) and –1 (for female). Since our classifier is based on linear regression, we are not restricted to binary codes. In principle, the vector r can contain any score encoding any attribute. Scores can be derived from other information directly associated with the walkers—for instance, their weight or their age. However, the scores can also be based on questionnaires completed by the subjects themselves, or by their physicians or therapists. This allows for a number of applications in biomechanics, clinical psychology, and neurology. For instance, it is well known that patients suffering from depression show walking patterns that differ from those of healthy controls (Lemke, Wendorff, Mieth, Buhl, & Linnemann, 2000). In an ongoing study we are currently quantifying these differences with our method, and we are developing tools to objectively assess the success of different therapies for depressive disorders (Michalak, Troje, Fischer, Heidenreich, & Schulte, in preparation).

In vision research, we are particularly interested in the perception of biological motion with respect to emotional attributes, personality traits, and other characteristics that we seem to be able to derive visually from the way a person moves. In an animation available at http://biomotionlab.ca/Demos/BMLwalker.html, we show examples of axes driven by the results of perceptual rating experiments. In addition to a "gender axis" and one that is based on the body mass index of the walker, we show two axes that were obtained by means of an experiment in which observers were shown a total of 80 different walkers displayed

as point-light displays on a computer monitor. Each display consisted of 15 white dots on a black background and was rendered from one of three different viewpoints (0 degrees = frontal view, 30 degrees, 90 degrees). A single rating session consisted of 80 trials, with each walker shown once for 7 s in a randomized order. All walkers within one session were shown from the same viewpoint. In order to indicate their rating, observers had to click one of six buttons displayed on the screen. Six observers participated in the experiments. For three of them the leftmost and rightmost buttons were labeled "nervous" and "relaxed," respectively. The other three observers were presented with the labels "happy" and "sad." Each observer carried out three sessions, one for each viewpoint, with short breaks between the sessions. The order of the three sessions was counterbalanced across observers.

The average of the ratings (across the three observers in each group and across the three different viewpoints) was used to form a vector r which in turn was used to compute the respective discriminant function v_d according to Equations 7 and 8. Animations along both the happy–sad axis and the nervous–relaxed axis give a clear percept of a change in the respective emotions of the walker. Visual inspection of the exaggerated walkers as well as quantitative examination of the discriminant function reveals the features that carry information about these attributes. Many of the differences between the nervous and the relaxed walker are due to the average pose and structure: walkers are perceived to be nervous when they have a skinny appearance with narrow pelvis and shoulders and when their shoulders and arms are pulled up tightly, whereas they are rated to be relaxed if they have wider frames and lower shoulders. With respect to their kinematics there is a shift of power for the horizontal movement of the markers, which is almost exclusively carried by the fundamental frequency in the relaxed walkers, to an increasing contribution of the second harmonic in the nervous walkers. Comparison of walkers along the happy–sad axis and inspection of the corresponding discriminant function shows that the main difference between walkers perceived to be sad or happy is the contribution of the second harmonic to the vertical movements. Here, the power of the second harmonic is relatively low for the sad walkers, while it is responsible for the appearance of bounciness in the happy walkers.

The data reported here are based on the relatively low number of only three observers per attribute. In fact, the power of the proposed

method is so strong that it produces reasonable results from very short experiments run with single, individual observers. We therefore designed a Web-based system in which users can generate their own axis based on ratings of a set of walkers (http://biomotionlab.ca/Demos/BMLrating). Upon entering the system, users are presented with an input mask that requests a few personal data (age, sex, country of origin) and then asks them to input an attribute of their choice along with two labels indicating the two ends of a Likert rating scale. For instance, users might input the attribute "sex" and then the labels "male" and "female" for the two ends of the scale. However, they are free to choose any attribute and any labels. Once this is done, the user will be presented with individual point-light walkers on half of the screen and a Likert scale with six buttons on the other half of the screen. The whole display will be titled with the attribute the user chose (e.g., "sex") and the first and last buttons contain the chosen labels (Fig. 12.3; also see color insert). After rating at least 20 walkers (but being encouraged to complete many more ratings), the user clicks a "finish" button; the system will then compute and display a discriminant function based on the obtained ratings. A point-light walker is displayed along with a slider that allows the viewer to interactively change the position of the walker on the axis (α in Eq. 9).

The success in revealing an axis that really reflects the intended attribute depends on how it gets mapped into our motion space. Not all attributes are expected to be represented linearly. For instance, consider a case in which the user chooses the attribute "symmetry," labeling the two ends of the Likert scale with "very asymmetric" and "very symmetric." He or she would probably attribute a high value ("very symmetric") to walkers that are very close to the average walker v_0 (Eq. 4). A walker with a strong asymmetry—for instance, the right arm swinging with a much larger amplitude than the left arm—would probably be at some distance from the average walker and would be assigned a low symmetry rating. However, a walker that is as asymmetric but with the left arm swinging more than the right arm would be located as far away from the average walker as the first asymmetric walker, but in the opposite direction, and would be assigned with the same rating. Fluctuating asymmetry is not distributed linearly in our space but rather concentrically, and any attempt to capture it with linear regression will fail.

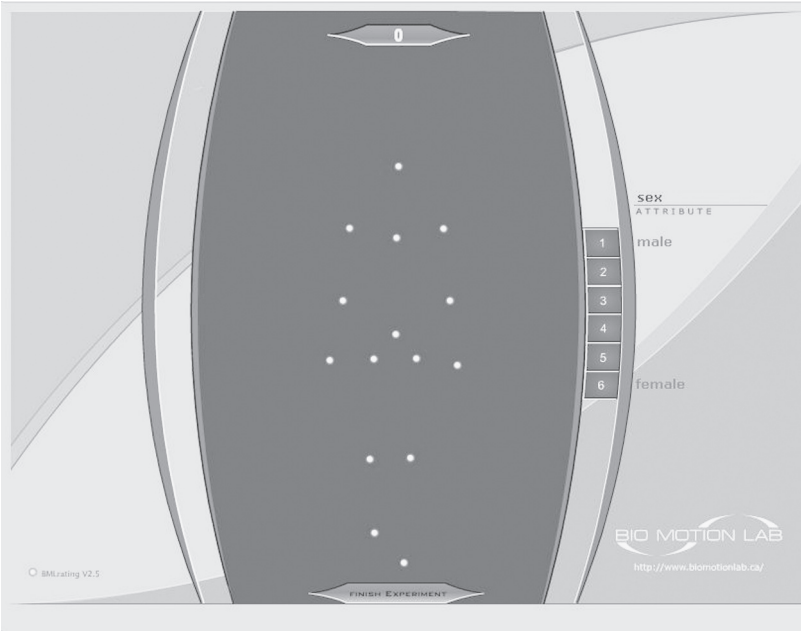

FIGURE 12.3. Layout of a Web-based demonstration. Observers can choose any attribute (e.g., "sex") along with two labels for the beginning and the end of a scale (e.g., "male" and "female") and then have to apply ratings to a series of individual point-light walkers. At the end they are presented with an animation that reflects their ratings in terms of a linear discriminant function.

Given this last consideration, it is surprising how many attributes are being successfully represented. While most users of our Web-based demonstration try attributes such as "sex," "attractiveness," "confidence," "mood," "weight," "age," "strength," "sportiness," etc., some came up with very creative ideas. For instance, we saw observers rate walkers according to their voting behavior (with labels "conservative" versus "liberal"), with very consistent outcomes, the discussion of which is beyond the scope of this chapter. One of my favorite axes was created by a user from Munich who rated walkers according to whether he would expect to see them in the rich and trendy neighborhood around Munich's Isartor or rather in a particularly shabby and rundown neighborhood characterized by cheap bars and plenty of nightlife.

We use this system t o demonstrate the richness of information conveyed by human motion to the public visiting our Web site, as well as to

students in lab courses. In addition, however, we use it to obtain ratings for a large number of trait terms, which eventually can be applied to factor-analytic procedures. This will help us to understand the topology and extract the cardinal dimensions of the perceptual space spanned by biological motion walking patterns.

Conclusion

Biological motion perception involves a complex hierarchy of visual information processing. A particularly interesting level is the one referred to as style recognition. Once an actor and the performed action are recognized on a basic level, style recognition can potentially reveal information about the specifics of an actor's identity, personality, and emotions. The framework we outlined in this chapter serves mainly two different functions. On the one hand, it helps us understand the complexity of the stimulus itself—a stimulus that is handled so effortlessly by our visual system. Understanding how our visual system solves the sophisticated problems involved in style recognition requires a comprehensive understanding of the constraints contained in the statistics of movement data and the encoding schemes for information in biological motion. Here, we approached the question of information encoding and retrieval from a pattern-recognition perspective. While we learned plenty about the way information is encoded in biological motion, the particular way it is retrieved by the human visual system may be very different from the way we did it here.

On the other hand, however, if considered as a model for information processing in the human visual system, our approach creates a number of hypotheses about its functioning that can well be tested. For instance, an artificial walker located 6 std away from the average walker on the sex axis is perceived completely unambiguously as male or female, even though such a walker probably has never been seen before in reality. Apparently, linear extrapolations in the proposed space result in walkers that are perceived as caricatures representing certain attributes even better than the real walkers—a strong argument for the idea that our visual system operates with similar representations. An item analysis with a close comparison between the artificial classification of individual walkers and the psychophysically obtained ratings also reveals striking

similarities. The same walkers that are misclassified by the linear system tend to be misclassified by human observers, and this also argues for similar representational spaces and metrics within these spaces.

While we are working here with human walking patterns, the framework described here can be extended to other movements as well. Each class of movements, however, requires its own description. A model for running could be obtained similarly to the way we obtained the walking model. However, at least within the framework outlined here, it would not make sense to try to describe both walking and running patterns within the same model. Our model is based on morphability. Each item in the space must match any other item in a canonical, unambiguous way. The correspondence between two items defines the "morph" between them (i.e., a smooth transition from one item to the other). Of course, it is possible to blend a walking pattern into a running pattern, but the blending is not unique, since the correspondence between the two patterns can be defined in several different ways. Dynamic models of gait production (Alexander, 1989; Golubitsky, Stewart, Buono, & Collins, 1998, 1999) show that the transition between walking and running is characterized by a singularity, and therefore these gaits represent two principally different motion patterns.

Similarly to other implementations of morphable models, our framework relies on establishing correspondence between features across the data set, resulting in a separation of the overall information into range-specific information on the one hand and domain-specific information on the other hand (Ramsay & Silverman, 1997). Applied to the current model, the range-specific information is the positional information contained both in the average pose and in the Eigenposes. The domain-specific information is the information about when things are happening. This information is contained in the fundamental frequency and in the phase of the walk. The domain-specific (i.e., temporal) part of the walking information therefore has a comparatively simple description.

For nonperiodic motions, a more complex formulation has to be employed. A very explicit way to do this is to define the temporal behavior of a motion in terms of the deviations with respect to a prototype. The prototype can be any typical example of the respective motion pattern. The temporal behavior of any other item can then be formulated explicitly in terms of the time warp required to minimize

the distance between the prototype and the time-warped version of the item. Implementations of such models have been described by Giese and Poggio (2000) and Ramsay and Silverman (1997). They could be incorporated in the framework described here in order to isolate information carrying biologically or psychologically relevant traits from actions that require more complex temporal descriptions, and to use it in turn to attribute personality and emotion to digital characters.

A number of modifications to the proposed model might further improve its value in classifying stylistic attributes and in identifying the features they are based on. Our approach is based on the attempt to linearize the data—that is, to transform it into a representation that describes it in terms of a low-dimensional convex linear manifold. Once this is achieved, we classify them by means of simple linear regression. Both the representation that we use as well as the classifier operating on it might not be optimal. For instance, there exist nonlinear methods for dimensionality reduction (e.g., Roweis & Saul, 2000) that might lead to latent variable representations that eventually result in a closer approximation to linearity than the combination of Fourier decomposition in the pose space and PCA in the walker space. There is also plenty of potential for improving the classifier. Linear regression is very sensitive to outliers. Even if we were assuming that the optimal classifier is eventually linear, methods like linear support vector machines (Cristianini & Shawe-Taylor, 2000) or robust PCA (De la Torre & Black, 2001) might improve classification. Nonlinear methods offer even more options.

By using these nonlinear methods, however, we might also have to sacrifice the ability to use the very same framework for the analysis of motion data on the one hand and synthesis and visualization on the other, which turn out to be of great help in combining pattern recognition approaches to biological motion perception with psychophysical methods. This combination, however, provides the basis for a mapping between a well-controlled, parameterized stimulus space and the perceptual spaces that biological motion research aims to explore.

References

Alexander, R. M. (1989). Optimization and gaits in the locomotion of vertebrates. *Physiological Reviews*, 69(4), 1199–1227.

Barclay, C. D., Cutting, J. E., & Kozlowski, L. T. (1978). Temporal and spatial factors in gait perception that influence gender recognition. *Perception & Psychophysics*, 23, 145–152.

Beauchamp, M. S. (2005). See me, hear me, touch me: Multisensory integration in lateral occipital-temporal cortex. *Current Opinion in Neurobiology*, 15(2), 145–153.

Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2002). Parallel visual motion processing streams for manipulable objects and human movements. *Neuron*, 34(1), 149–159.

Bertenthal, B., Proffitt, D. R., & Kramer, S. J. (1987). Perception of biological motion by infants: Implementation of various processing constraints. *Journal of Experimental Psychology: Human Perception and Performance*, 13, 577–585.

Bidet-Caulet, A., Voisin, J., Bertrand, O., & Fonlupt, P. (2005). Listening to a walking human activates the temporal biological motion area. *Neuroimage*, 28(1), 132–139.

Blake, R. (1993). Cats perceive biological motion. *Psychological Science*, 4(1), 54–57.

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(11), 3737–3744.

Bruderlin, A., & Williams, L. (1995). Motion signal processing. Paper presented at the Proceedings of SIGGRAPH 95, Los Angeles.

Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., et al. (2004). Neural circuits involved in the recognition of actions performed by nonconspecifics: A*n FMRI study. Journal of Cognitive Neuroscience*, 16(1), 114–126.

Cavanagh, P., Labianca, A. T., & Thornton, I. M. (2001). Attention-based visual routines: Sprites. *Cognition*, 80(1-2), 47–60.

Cristianini, N., & Shawe-Taylor, J. (2000). An introduction to support vector machines and other kernel-based learning methods. Cambridge: Cambridge University Press.

Cutting, J. E. (1978). Generation of synthetic male and female walkers through manipulation of a biomechanical invariant. *Perception*, 7(4), 393–405.

Cutting, J. E., & Kozlowski, L. T. (1977). Recognizing friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomic Society*, 9(5), 353–356.

Cutting, J. E., Moore, C., & Morrison, R. (1988). Masking the motions of human gait. *Perception & Psychophysics*, 44(4), 339–347.

Cutting, J. E., Proffitt, D. R., & Kozlowski, L. T. (1978). A biomechanical invariant of gait perception. Journal of Experimental Psychology: *Human Perception & Performance*, 4, 357–372.

Davis, R. B., Ounpuu, S., Tyburski, D., & Gage, J. R. (1991). A gait analysis data collection and reduction technique. *Human Movement Science*, 10, 575–587.

De la Torre, F., & Black, M. J. (2001). Robust principal component analysis for computer vision. Paper presented at the ICCV-2001, Vancouver.

Dittrich, W. H., Lea, S. E. G., Barrett, J., & Gurr, P. R. (1998). Categorization of natural movements by pigeons: Visual concept discrimination and biological motion. Journal of the Experimental Analysis of Behaviour, 70, 281–299.

Dittrich, W. H., Troscianko, T., Lea, S. E. G., & Morgan, D. (1996). Perception of emotion from dynamic point-light displays represented in dance. *Perception*, 25(6), 727–738.

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, 293(5539), 2470–2473.

Farah, M. J., Tanaka, J. W., & Drain, H. M. (1995). What causes the face inversion effect? Journal of Experimental Psychology: *Human Perception and Performance*, 21(3), 628–634.

Fox, R., & McDaniel, C. (1982). The perception of biological motion by human infants. *Science*, 218(4571), 486–487.

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.

Golubitsky, M., Stewart, I., Buono, P. L., & Collins, J. J. (1999). Symmetry in locomotor central pattern generators and animal gaits. *Nature*, 401(6754), 693–695.

Golubitsky, M., Stewart, I., Buono, P.-L., & Collins, J. J. (1998). A modular network for legged locomotion. *Physica D*, 115, 56–72.

Grezes, J., Costes, N., & Decety, J. (1998). Top-down effect of strategy on the perception of human biological motion: A PET investigation. Cognitive *Neuropsychology*, 15(6–8), 553–582.

Grossman, E. D., & Blake, R. (2002). Brain areas active during visual perception of biological motion. *Neuron*, 35(6), 1167–1175.

Grossman, E. D., Blake, R., & Kim, C. Y. (2004). Learning to see biological motion: Brain activity parallels behavior. *Journal of Cognitive Neuroscience*, 16(9), 1669–1679.

Grossman, E. D., 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(5), 711–720.

Guo, S., & Roberge, J. (1996). A high-level control mechanism for human locomotion based on parametric frame space interpolation. In R. Boulic & G. Hegron (Eds.), Computer animation and simulation '96 (pp. 95–107). New York: Springer.

Ikeda, H., Blake, R., & Watanabe, K. (2005). Eccentric perception of biological motion is unscalably poor. *Vision Research*, 45(15), 1935–1943.

Jastorff, J., Kourtzi, Z., & Giese, M. A. (2006). Learning to discriminate complex movements: Biological versus artificial trajectories. *Journal of Vision*, 6, 791–804.

Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception & Psychophysics*, 14(2), 201–211.

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

Jones, M. J., & Poggio, T. (1999). Multidimensional morphable models: A framework for representing and matching object classes. *International Journal of Computer Vision*, 29, 107–131.

Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex of a walker from a dynamic point-light display. *Perception & Psychophysics*, 21(6), 575–580.

Kozlowski, L. T., & Cutting, J. E. (1978). Recognizing the sex of a walker from a dynamic point-light display: Some second thoughts. *Perception & Psychophysics*, 23(5), 459.

Lemke, M. R., Wendorff, T., Mieth, B., Buhl, K., & Linnemann, M. (2000). Spatiotemporal gait patterns during over ground locomotion in major depression compared with healthy controls. *Journal of Psychiatric Research*, 34(4–5), 277–283.

Marey, E.-J. (1895/1972). Movement. New York: Arno Press.

Mather, G., & Murdoch, L. (1994). Gender discrimination in biological motion displays based on dynamic cues. *Proceedings of the Royal Society of London Series B*, 258, 273–279.

Michalak, J., Troje, N. F., Fischer, J., Heidenreich, T., & Schulte, D. (in preparation). The embodiment of depression: *Gait patterns of currently and formerly depressed patients*.

Muybridge, E. (1887/1979). Muybridge's complete human and animal locomotion. New York: Dover.

Oram, M. W., & Perrett, D. I. (1994). Responses of anterior superior temporal polysensory (STPa) neurons to "biological motion" stimuli. *Journal of Cognitive Neuroscience*, 6(2), 99–116.

Peelen, M. V., & Downing, P. E. (2005). Selectivity for the human body in the fusiform gyrus. *Journal of Neurophysiology*, 93(1), 603–608.

Peuskens, H., Vanrie, J., Verfaillie, K., & Orban, G. A. (2005). Specificity of regions processing biological motion. *European Journal of Neuroscience*, 21(10), 2864–2875.

Pollick, F. E., Paterson, H. M., Bruderlin, A., & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition*, 82(2), B51–B61.

Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. Journal of *Neuroscience*, 18(6), 2188–2199.

Ramsay, J. O., & Silverman, B. W. (1997). Functional data analysis. New York: Springer.

Rosch, E. (1988). Principles of categorization. In A. Collins & E. E. Smith (Eds.), Readings in cognitive science: A perspective from psychology and artificial intelligence (pp. 312–322). San Mateo, CA: Morgan Kaufmann.

Rose, C., Bodenheimer, B., & Cohen, M. F. (1998). Verbs and adverbs: Multidimensional motion interpolation using radial basis functions. *Computer Graphics and Applications*, 18(5), 32–40.

Roweis, S. T., & Saul, L. K. (2000). Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290(5500), 2323–2326.

Shelton, C. R. (2000). Morphable surface models. International Journal of *Computer Vision*, 38, 75–91.

Thornton, I. M., Rensink, R. A., & Shiffrar, M. (2002). Active versus passive processing of biological motion. *Perception*, 31(7), 837–853.

Troje, N. F. (2002a). Decomposing biological motion: *A framework for analysis and synthesis of human gait patterns. Journal of Vision*, 2(5), 371–387.

Troje, N. F. (2002b). The little difference: Fourier-based synthesis of gender-specific biological motion. In R. Würtz & M. Lappe (Eds.), *Dynamic Perception* (pp. 115–120). Berlin: Aka Press.

Troje, N. F., & Vetter, T. (1998). Representations of human faces. In C. Taddei-Ferretti & C. Musio (Eds.), *Downward processing in the perception representation mechanism* (pp. 189–205). Singapore: World Scientific.

Troje, N. F., & Westhoff, C. (2006). The inversion effect in biological motion perception: Evidence for a "life detector"? *Current Biology*, 16(8), 821–824.

Troje, N. F., Westhoff, C., & Lavrov, M. (2005). Person identification from biological motion: *Effects of structural and kinematic cues. Perceptual Psychophysics*, 67(4), 667–675.

Unuma, M., Anjyo, K., & Takeuchi, R. (1995). Fourier principles for emotion-based human figure animation. Paper presented at the Proceedings of SIGGRAPH 95.

Urtasun, R., Glardon, P., Ronan, B., Thalmann, D., & Fua, P. (2004). Style-based motion synthesis. *Computer Graphics*, 23, 799–812.

Vaina, L. M., Solomon, J., Chowdhury, S., Sinha, P., & Belliveau, J. W. (2001). Functional neuroanatomy of biological motion perception in humans. *Proceedings of the National Academy of Sciences U S A*, 98(20), 11656–11661.

Vallortigara, G., & Regolin, L. (2006). Gravity bias in the interpretation of biological motion by inexperienced chicks. *Current Biology*, 16(8), R279–R280.

Vallortigara, G., Regolin, L., & Marconato, F. (2005). Visually inexperienced chicks exhibit spontaneous preference for biological motion patterns. *PLoS Biology*, 3(7), e208.

Vetter, T., & Poggio, T. (1997). Linear object classes and image synthesis from a single example image. IEEE Transactions on Pattern Analysis and Machine Intelligence, 19, 733–742.

Vetter, T., & Troje, N. F. (1997). Separation of texture and shape in images of faces for image coding and synthesis. Journal of the Optical Society of America A, 14, 2152–2161.

Witkin, A., & Popovic, Z. (1995). Motion warping. Computer Graphics Proceedings of SIGGRAPH 95, 105–108.