

What is biological motion?: Definition, stimuli and paradigms

Nikolaus F. Troje

Department of Psychology,
Department of Biology,
School of Computing,
and Centre for Neuroscience Studies,
Queen's University, Kingston, Canada

to be published in:

Rutherford, M. D. and Kuhlmeier, V. A. (eds.) *Social Perception: Detection and Interpretation of Animacy, Agency, and Intention*. MIT Press, 2013

Gunnar Johansson's biological motion: structure from non-rigid motion

The term “biological motion” is closely associated with the work of Gunnar Johansson (1911 – 1998), an experimental psychologist who received his PhD from the University of Stockholm and then taught and conducted research at the University of Uppsala, Sweden, for most of his career. Hardly any study on biological motion is published without a reference to at least one of the two papers which Johansson published on his observation that an observers effortlessly make out a human figure behind the degraded visual stimulus that consists of nothing but a few dots moving along with the major joints of a human body (Johansson, 1973, 1976).

Gunnar Johansson was an extremely prolific researcher, well networked, and with strong ties to many of the personalities that had shaped European post-war experimental psychology. In addition to the two papers on the perception of human motion, he published almost 100 papers over the course of his career. He was strongly influenced by the school of Gestalt Psychology. Wolfgang Köhler had visited him several times in Uppsala and the two had planned collaborative work shortly before Köhler's unexpected death in 1967. Johansson also maintained a long academic relationship with James J. Gibson. With Gibson he shared the interest in the question how the visual system samples and then extracts from the ambient optic array information about discrete objects and events – the ambient optic array being the ever changing continuous, yet structured stream of light irradiating the location of the observer from all directions. Particularly, the two friends discussed (and sometimes argued about) the way the visual system distinguishes between the visual motion induced by the observer himself while moving through the environment, and the motion of objects, animals, and people relative to the visual environment.

Johansson's approach to this question was his “vector analysis” theory which he had first laid out in his doctoral thesis (Johansson, 1950). The theory was rooted in Gestalt theory, specifically in Duncker's work on induced motion (Duncker, 1928), and it was based on Johansson's own experimental work for his doctoral thesis under the supervision of Gestalt psychologist David Katz at Stockholm University. Johansson's experiments had set out to demonstrate how the human visual system would break down visual motion of simple objects into “common” motion encoded in a global reference system, and “relative” motion represented within an object-based reference system. Common motion of the entire visual field with respect to the observer would be indicative for observer motion, while motion relative to the global, allocentric reference system has the potential to characterize objects – both in terms of motion parallax induced by the observer but also by indicating the object's own motion. On the object level itself, the same distinction seemed to make sense, too. The motion common to all the parts of an object describes its trajectory through space, while the relative motion between the parts bears information about changes of orientation or the configuration of the internal structure of a deformable object. “Vector analysis” emphasized on the idea that the visual system would decompose motion in the optic array into components of motion occurring within a hierarchy of nested reference systems (Johansson, 1974).

What Johansson later called “biological motion” seemed to be an ideal domain to further explore these ideas and provided a natural continuation of his earlier studies. “The biological motion effect is a most efficient demonstration and verification of the theory of perceptual vector analysis”, Johansson said towards the end of his career in an interview conducted by William Epstein in 1992 (Jansson, Bergstroem, & Epstein, 1994). To Johansson, the point-light display technique and the term “biological

motion” that he used for it in his first paper on the topic (Johansson, 1973) therefore served a very specific purpose. He was interested in the mechanisms of perceptual organization. Specifically, he was interested in the strategies the visual system is using to integrate the individual moving dots into the percept of the coherent, articulated shape of a human or animal body. He used his new stimuli to demonstrate that the visual system decomposes motion not just into common and relative motion, but into a whole hierarchy of partial motions, each of which encoded relative to its parent coordinate system. In a way, his work and theories laid the grounds for the interest into visual “structure-from-motion” which soon became a popular research topic within vision science (e.g. Ullman, 1979). A number of early papers in biological motion research elaborated on Johansson's general idea and provided computational methods to retrieve structure from the non-rigid motion of a moving body (Cutting, 1981; Hoffman & Flinchbaugh, 1982; Webb & Aggarwal, 1982).

Today, research on “biological motion” perception has deviated in many ways from the questions which motivated Johansson in the first place. In fact, in the above mentioned interview (Jansson, et al., 1994), Johansson expressed some degree of frustration about referencing to his work in contexts which do not reflect his initial interest in the perceptual dissociation between common and relative motion. However, Johansson, in his early work, did not explicitly define what he meant by the term “biological motion” and how he wanted it to be used. Rather, than introducing “biological motion” as a technical term, he used it in an informal way to refer to the stimulus domain from which he derived the displays that he needed to further his studies on the role of motion in perceptual organization: Stylized depictions of the articulated motion of the body of humans or other animals. Biological motion in this sense is obviously a rich source of information about conspecifics, as well as non-conspecific animals which may serve as prey or provide a predatory threat to the observer. As such, biological motion in this broader sense provides a very interesting stimulus class for the vision researcher. Particularly for social animals such as humans, the rich information about the identity of an agent as well as its actions and intentions that is contained in the way an agent moves affords adaptations to exploit it. Similar information may also be carried by other information channels, for instance facial features, vocalizations, or body posture. However, motion as a source for socially relevant information might be the evolutionarily oldest resource and certainly still plays a key role in person perception in humans. As such, it motivates and justifies the wide interest which biological motion research has experienced particularly over the last 15 years.

As interesting as it is, these newer avenues have little to do with what Gunnar Johansson was initially concerned with. The connections between Johansson's work and the emerging field of social perception are only superficial. One such connection is that point-light displays seemed to provide a good solution to a technical problem. Studies on social perception from movement patterns often attempt to isolate motion from other sources of information, and point-light displays certainly provide that to a reasonable extent. Specifically, they are deprived of facial information, hair, clothing, and at least some aspects of the morphology of the body. Note, however, that information about the general proportions of the body is still preserved. The motion of the body is still carried by a substrate that in itself contains some information about the actor and action.

The other connection might be the term itself. “Biological motion” has become tightly associated with point-light displays of Johansson's work. In a more colloquial sense, however, it has a much wider meaning. What other label but “biological motion” could we use when we talk about the movements of animals and humans and the information they convey? “Animate motion” (not to be confused with “animated motion”) would be a good term if we wanted to avoid the term “biological motion” and its association to Johansson's question about structure from motion. However, the term “animate motion”

suffers a somewhat similar fate as “biological motion”. It has been adopted by another interesting domain within the field of social perception, a domain, which is in many respects complementary to what is commonly associated with biological motion. Biological motion covers intrinsic motion, that is, the motion of a deforming body relative to reference systems centred on the body or its parts. The term “animate motion”, on the other hand, gets commonly associated with the motion of whole objects through space and relative to each other. In order to isolate this aspect, rigid shapes such as triangles, circles, or squares are often used to study animate motion. Under certain conditions, such rigid object movement is readily interpreted by a human observer as intentional and thus animate, as was first demonstrated by Heider and Simmel (1944). For a review on more recent studies on animate motion, see Scholl (2000) and chapter ?? in this book. With respect to Johansson's initial terminology, biological motion is primarily concerned with “relative” motion while animate motion is all about the “common motion” of the entire object through space.

With both “biological motion” and “animate motion” being already reserved for more specific aspects, I will use in this chapter the term “life motion” for visual motion that expresses any sort of aspect characteristic for the motion of living beings. “Life motion” includes any kind of stimulus that elicits the percept of something or someone being alive, no matter what specific depiction is being used and what aspect of the motion is being emphasized. I am aware that the term “life motion” is at risk to suffer a similar fate as “biological motion” and “animate motion” as it has recently been used to label a very specific aspect contained in the local motion of distal body parts (e.g. Wang, Zhang, He, & Jiang, 2010). At this point, however, it appears still a reasonable term to be used for the most general of the potential interpretations of biological motion.

This allows me – at least for the purpose of the current chapter – to reserve the term “biological motion” for stimuli and studies concerned with the intrinsic, non-rigid motion of the deforming body, rather than the motion path that its centre of gravity subscribes through space. This usage includes Johansson's question about the retrieval of articulated structure, but does not restrict it to this topic. Staying in line with the predominant way the term was used in the literature, it also includes a wealth of work interested in what the visual system can retrieve from the motion of animals and people once the structure-from-motion problem is solved. Also note, that this definition doesn't restrict us to point-light displays. Any stimulus that attempts to isolate the intrinsic motion component from other sources of information about actor and action is suitable to study biological motion. If not specifically interested in the structure-from-motion problem, stick figures may be as good or even better than point-light displays to study the role of intrinsic motion without the influence of facial features or detailed information about the morphology of a body. The amount of additional information and the degree of realism of a depiction, however, is an important dimension when talking about stimuli used in life motion research and will be further discussed at the end of this chapter.

Experimental paradigms: detection and direction

Restricting the term “biological motion” to the study of the perception of the “relative”, intrinsic deformation of a body in action still encompasses a large variety of questions and methods. In an attempt to further investigate what is meant by “biological motion perception” I will try to categorize and critically discuss the different experimental paradigms and perceptual tasks that investigators have used to assess biological motion perception in human observers.

In trying to bring some order to the large host of literature, a reasonable distinction is the one between studies that aim to assess biological motion perception in general, and the ones that ask very specific questions about what kind of information can be retrieved from biological motion stimuli.

Assessing biological motion perception in general might be interesting with respect to different context conditions or different observer populations. For instance, after having demonstrated the general biological motion phenomenon, Johansson varied display times in order to measure how much time it takes to perceptually organize the dots into the shape of a person (Johansson, 1976). Subjects were shown point-light displays of a walker in sagittal view at different exposure times and were then asked to verbally report what they saw. They generally reported seeing a human walker and only for exposure times shorter than 200 ms did their performance decrease. James Cutting, who was one of the first researchers to advance Johansson's early work, set out to measure the robustness and salience of biological motion point-light displays presented in visual noise (Cutting, Moore, & Morrison, 1988). He showed stationary (as if on a treadmill) point-light walkers from a sagittal view and asked his observers to indicate whether they were facing left or right while manipulating the nature of the mask. He basically showed that a so-called “scrambled mask”, that is, a mask made of dots that move in the same way as the dots that constitute the walker, is much more efficient than any other kind of dot mask. Berthenthal and Pinto (1994) used a similar stimulus, but a different task to investigate the role of global structure versus local features for biological motion perception: They also showed a mask made of scrambled walkers and embedded a coherent walker into it. Only half of the trials really contained the coherent walker while the other half showed the mask alone. The observer's task was to identify which of two displays contained the walker. Comparing responses to upright and inverted versions of this detection task, and finding a significant inversion effect, it was concluded that biological motion perception is driven by global, “configural” processes and does not require the prior detection of individual features or local relations. This conclusion was based on earlier findings that the face inversion effect is mainly caused by orientation sensitivity of configural, but not featural processing.

These three examples are based on three different tasks. Johansson's demonstration employed a free response format in which observers reported their perception of an unmasked point-light display. Performance was in most cases at ceiling, demonstrating the strength of the basic biological motion phenomenon. Masking, as used in Cuttings (1988) and also in Berthenthal and Pinto's (1994) study helps to reduce performance and brings it into a range where we can expect to sensitively measure the effects of other experimental manipulations. But it also introduces a confound: Observers might perform poorly because they have problems with the required figure-ground segregation, or they might perform poorly because they have difficulties with the perceptual organization required to turn the individual dots into the coherent, articulated shape of a human body. The direction discrimination task used by Cutting adds another uncertainty: After having segregated the target walker from the masking dots, observers then have to determine the facing direction of the walker from whatever cues are available.

It turns out there are a few of them – at least more than one: An important cue is certainly the motion-mediated structure of the body: Even from a single frame, it would probably be easy to determine facing direction if the articulated structure of the body was readily available – for instance by drawing explicit connections between the dots, thus turning the point-light display into a stick figure. Motion might simply help to replace the missing sticks. The cues to direction are then contained in the reconstructed posture of the body.

The efficiency of a mask made of scrambled walkers already points to a second source of information. It turns out, in fact, that the local motion of the individual dots can also carry sufficient information about facing direction. I will elaborate on that mechanism further below.

The direction task can thus be solved by the observer in two entirely independent ways. Problems with one of them might well be compensated by the other and variance in performance between conditions and observer groups are therefore hard to interpret. Along with the ability to segregate the foreground

walker from the background mask, three different perceptual abilities are therefore confounded in the popular biological motion direction task.

The experimental paradigms discussed above have all been used with special participant populations to assess their ability to perceive biological motion. Several modifications exist and new ones have been introduced. For instance, Mather and colleagues (Mather, Radford, & West, 1992) used a task in which coherent sagittally presented walkers were intermixed with composite walkers in which upper and lower body were facing into opposite directions. Observers had to detect the incoherence. Gurnsey and colleagues (2010) introduced a modified version of the direction task: Rather than using sagittal views, they reduced the angle between the left-facing and right-facing views to threshold levels – thus avoiding the need to use a mask and therefore the confound with figure-ground segregation. Ahlstrom and colleagues (1997) used a modification of the detection task described above in which they employed not just point-light walkers but depictions of all sorts of other actions, too. This particular stimulus set is important because it has later become the basis for identifying and localizing the superior-temporal sulcus (STS) as a primary area for the processing of biological motion (Grossman et al., 2000).

Beyond structure from motion: style

While these experimental paradigms have been used for general assessment of an observer's ability to derive the articulated structure and kinematics of a point-light depiction, other tasks have been used to further explore the specific nature of the information that can be derived from such displays. The particular style of the movement can reveal properties of the actor that go beyond its mere detection and orientation. Biological motion point-light displays have been used to explore the limits of this ability and to investigate which particular attributes are carrying the information. James Cutting and colleagues were already able to show that observers can derive the sex of a walker (Kozlowski & Cutting, 1977) from point-light displays and that they were even able to recognize familiar individuals (Cutting & Kozlowski, 1977). A number of studies later used principled manipulations of the point-light displays to investigate the role of motion-mediated structure, on the one hand, and the kinematics of the body, on the other hand (Mather & Murdoch, 1994; Troje, 2002; Troje, Westhoff, & Lavrov, 2005; Westhoff & Troje, 2007) for gender classification and person identification. Both domains definitely play a role but information encoded in the kinematics of a walker seem to dominate the information obtained from structural cues.

Other properties encoded in the particular style of walking and other movements involve age (Montepare & Zebrowitz-McArthur, 1988), emotional attributes (Atkinson, Dittrich, Gemmell, & Young, 2004; Clarke, Bradshaw, Field, Hampson, & Rose, 2005; Dittrich, Troscianko, Lea, & Morgan, 1996), and personality traits (Troje, 2008). The kinematics of the point-light animations can also reveal sophisticated details about the performed action itself, such as the weight of a lifted box (Runeson & Frykholm, 1981), and the distance of a thrown object (Knoblich & Flach, 2001; Munzert, Hohmann, & Hossner, 2010). Even deception (Runeson & Frykholm, 1983) and intention (Hohmann, Troje, Olmos, & Munzert, 2011) can be derived from stylistic properties of human motion as encoded in point-light displays.

Note, that while all these studies fall into the domain of biological motion perception and often employ point-light displays, they have little in common with the initial question about perceptual organization of the individual dots into a coherent figure. Deriving sophisticated information about actor and action from the stylistic details of a movement most likely requires access to the articulated structure of the body, but how this structure is obtained might not be relevant at all. If it is relevant, it might rather

create a confounding variable which hinders the interpretation of the resulting data. Point-light displays are often used because they seem to isolate the kinematics of a movement from all sorts of other sources of information present in more realistic depictions. However, they are not entirely free of structural information which can be reconstructed relatively easily, as Johansson and the early biological motion researchers have shown. If the subject of an investigation is not the structural reconstruction process itself, but rather the semantic contents of movement of the articulated figure it may be better to present the articulation explicitly to the observer and replace the point-light walker with a stick figure or other representations that provide the structure of the body directly while still depriving the viewer from morphological details and facial information.

This approach has rarely been taken within visual psychology, but is employed regularly by researchers in computer animation (e.g. Brand & Hertzmann, 2000; Calvert, 2007; Hertzmann, O'Sullivan, & Perlin, 2009; Hodgins, O'Brien, & Tumblin, 1998; Unuma & Takeuchi, 1993). Even though this literature did not grow out of the tradition of Johansson's introduction of "biological motion" point-light displays and does not use the same terminology, it is concerned with the same questions we discussed in the previous paragraphs: How is information about identity, emotion and personality encoded in the stylistic details of an actors movements? In computer animation, this question is directly related to a very applied one: What is required to generate human or animal motion that convincingly simulates natural human movement while retaining information about the actor? In the case of photorealistic animations in movies and games we seek an answer to the questions: Which are the sophisticated artifacts that reveal to the observer the artificial nature of the animation and how can they be avoided?

Before structure from motion: life detection

In addition to work on the structure-from-motion problem and the studies exploring stylistic contents, there is a third domain that adds to the general phenomenology and the study of biological motion perception. The fact that masking a walker with scrambled versions of the same walker still allows recognition to some degree has been used to argue for the importance of global processing (e.g. Bertenthal & Pinto, 1994). After all, the additional dots mask the local motion and performance must therefore be carried by global structure. However, the argument can be reversed. While observers can still make the walker out if the mask is not too dense, it is clear that the scrambled walker mask is extremely efficient when compared to other kinds of dot masks. For instance, Cutting (Cutting, et al., 1988) reports experiments in which 55 dots were sufficient to reduce direction discrimination performance to chance level if a scrambled mask was used. If the same number of noise dots moved either on linear or circular trajectories, observers responded still correct in about 80% of the cases (see also Hiris, Humphrey, & Stout, 2005). This shows that depriving a point-light stimulus of its local motion components seriously impairs recognition.

Kozlowski and Cutting (1977) were probably the first who suggested that the motion of individual dots may contain information entirely independent of their contribution to the retrieval of articulated shape. Their finding that the single dot representing a foot is enough for an observer to discriminate between a male and a female walker was later retracted (Kozlowski & Cutting, 1978), but others were able to confirm that the local motion of individual dots does help to interpret biological motion (Mather, et al., 1992). Troje and Westhoff (2006) used spatially and temporally scrambled versions of a biological motion point light walker to show that even if deprived of any structural information, observers can still determine the facing direction of a walker.

They also showed that the ability to do so is orientation dependent and breaks down entirely when the

scrambled walker is presented upside down. This inversion effect challenges the close association of inversion effects and “configural processing” which originated from the face recognition literature but had also been adopted in biological motion research (Bertenthal & Pinto, 1994; Farah, 1992). In addition, it provided a tool for the search of the source of the signatures that communicate facing direction. It turned out that inverting the trajectory of only the feet entirely disrupted the ability to obtain facing direction from scrambled biological motion (Troje & Westhoff, 2006) and that the critical property of that trajectory was the vertical acceleration pattern and whether or not it was consistent with gravitational acceleration (Chang & Troje, 2009a). Troje and Westhoff’s finding also prompted a discussion about a more general purpose of the visual filter that provides the sensitivity to this particular feature. It was suggested that it functions as a general “life detector” that labels a motion stimulus as potentially animate (Johnson, 2006). Evidence that sensitivity to the invariants in foot motion exists in non-human species (Vallortigara & Regolin, 2006) and that it might be present in human neonates prior to individual experience (Bardi, Regolin, & Simion, 2011; Simion, Regolin, & Bulf, 2008) points to an evolutionarily old origin. Furthermore, it was shown that the same features that convey information about facing direction also contribute to the perceived animacy of a point light stimulus (Chang & Troje, 2008). A number of recent studies have confirmed the idea that visual invariants contained in the way the feet respond to constraints imposed by gravity and inertial forces are in fact used by the visual system to direct attention and guide orientation responses to people and other terrestrial, legged, animals independently of their particular shape (Hirai, Saunders, & Troje, 2011; Jiang & He, 2007; Wang, et al., 2010) (for a review of the literature, see Troje & Chang, in press). A recent study showed that upright scrambled biological motion displays are perceived to be significantly longer in duration than inverted stimuli (Wang & Jiang, in press).

All these studies employed stimuli similar to the ones used in Troje and Westhoff’s (2006) initial study. As it turns out, the local cues can be even much stronger than implicated by their data. Hirai and colleagues (Hirai, Chang, Saunders, & Troje, 2011) recently described a point light stimulus which is entirely free of structural cues to facing direction, but still elicits facing direction performance at almost ceiling levels. Stimuli were derived from a normal sagittal view of a point-light walker by randomly displacing all dots as in the standard scrambled walker – but this time only disrupting the horizontal configuration while leaving the vertical organization intact. In a similar way as vertical acceleration “validates” motion as “biological” without being able to provide information about facing direction itself, the vertical location of the dots representing the feet (either above or below the rest of the body) turns out to be a very effective validating cue.

For our discussion of the meaning and usage of the term “biological motion” the ability to exploit purely local information has two important implications. The first one, I already mentioned briefly above: In general, the frequently used direction discrimination task (“Which side is the walker facing?”) can be solved using two very different and to a large degree independent sources of information, namely the motion-mediated articulated structure, on the one hand, and the specific, local motion of the feet on the other hand. Using this task to assess biological motion perception in specific populations or under specific conditions measures two very different abilities within a paradigm that confounds them such that it becomes very difficult to arrive at clear conclusions. The same is a potential problem in other paradigms, too. For instance, if the same local cues that indicate facing direction are used to label motion as biological as suggested above (Chang & Troje, 2008; Johnson, 2006; Troje & Westhoff, 2006) the local motion of the feet may as well be used in detection experiments similar to the one introduced by Bertenthal and Pinto (1994).

The second implication is the following: If we accept that the ability of an observer to exploit local

invariants to derive information about animacy and orientation of an agent falls into the domain of “biological motion perception”, then scrambled biological motion cannot be regarded as “non-biological motion”. That, however, is often the case. Scrambled biological motion is frequently used as a control stimulus in behavioural, physiological, neuropsychological, or imaging studies (e.g. Grossman, et al., 2000; Vaina, Solomon, Chowdhury, Sinha, & Belliveau, 2001). In principal there is no problem with this kind of contrast, as long as the goal, the limits, and the implications of such experiments are clear. Contrasting the perception of coherent and scrambled biological motion captures an observers ability to retrieve motion-mediated structure from a point light display. It does not, however, allow more general conclusions about biological motion perception. The same is true for the processing abilities of a brain area identified by means of fMRI responses to that contrast.

The question about which brain areas are involved in biological motion perception is even more complicated. A central area which keeps reappearing in the literature reporting brain imaging studies on biological motion perception is the superior temporal sulcus (STS). To which degree STS is really a “biological motion area” and what kind of aspects it processes is not clear at all. As mentioned already above, an experiment that effectively localizes STS involves the contrast between coherent point-light displays of actions other than walking and scrambled versions of the same displays. It therefore seems to respond to the structure-from-motion aspects of biological motion. The problem is that the response is neither specific nor does it generalize. On the one hand, STS responds not just to biological motion but also to many other visual and even non-visual social stimuli (Allison, Puce, & McCarthy, 2000; Kriegstein & Giraud, 2004; Puce & Perrett, 2003). On the other hand, the response to biological motion point-light displays seems to be specific to certain actions but doesn't generalize to others. For instance, STS does not necessarily respond to the stimulus most widely used in behavioural experiments on biological motion: the display of a walker shown in sagittal view as walking on a treadmill (Jastorff & Orban, 2009).

There are more examples that demonstrate how important it is to clearly state what an experiment is aiming at and what exactly is investigated when talking about biological motion. For instance, using a detection task Ikeda, Blake and Watanabe (2005) found that biological motion is particularly poor in the visual periphery – much worse than what could be explained by decreasing visual acuity. It was later discovered, however, that this phenomenon depended on the walker being masked by noise dots. Using a paradigm that did not require masking noise Gurnsey and colleagues (R. Gurnsey, Roddy, Ouhana, & Troje, 2008) were able to compensate the decreasing performance at greater eccentricities by increasing the size of the stimulus (see also B Thompson, Hansen, Hess, & Troje, 2007). Only detecting a walker in masking noise was ‘unscalably poor’, but not the ability to determine facing direction of a point-light walker.

Two other studies also provide data that should make us cautious about using masking noise to reduce ceiling performance and measure thresholds. In a study comparing older and younger adults in a direction discrimination task, performance was equal between the two groups except when masking noise was added, in which case the older adults were significantly worse than the younger adults (Pilz, Bennett, & Sekuler, 2010). Here, as well as in a study that compared amblyopic perception of biological motion to vision in normal controls (B. Thompson, Troje, Hansen, & Hess, 2008) the results suggest that the deficits observed in older or visually impaired participants were mainly due to a decreased ability to distinguish the walker from the noisy background – much more so, at least, than to structure-from-motion related aspects of biological motion perception or the ability to process local “life detection” invariants.

A taxonomy of “biological motion”

I tried to position the different usages of the term “biological motion” in a space that is spanned by the questions that motivate the research, the experimental paradigms being used, and the stimuli employed to depict human (or animal) motion. Figure 1 attempts to summarize some of the considerations made in the previous sections. I consider the scheme depicted here as work in progress and the basis for further discussion, rather than a final solution to the question I am asking in the title of this chapter.

The stimuli used to study visual life motion of humans and animals often are degraded and stylized to some degree in order to isolate the aspects of interest from confounding variables that are correlated with motion signals in more realistic stimuli (video, full computer animations) or the ones experienced in the real world. Some studies have chosen to use less abstracted depictions, such as more realistic shaded and/or textured 3D avatars, whose 3D morphological attributes are kept constant (e.g. Hodgins, et al., 1998). A potential problem with this approach is that morphological and kinematic parameters are highly correlated in real life motion, and that the human visual system is probably very sensitive to violations of learned correlations. Rather than studying kinematics in isolation, realistic but normalized morphologies might produce severe consistency effects which depend on how well a given motion pattern fits the chosen geometry of the normalized avatar.

As already discussed above, I suggest to restrict the term “biological motion” to the study of the perception of intrinsic, deformable motion of animals and people by means of stimuli that are stylized in order to keep information conveyed by other channels but the kinematics of the body at a minimum. The focus on intrinsic rather than extrinsic motion might feel somewhat arbitrary and unnecessary in light of questions surrounding cognitive mechanisms of social perception – which is a major focus in much of the more recent literature on biological motion perception. However, the distinction makes perfect sense when investigating the perceptual mechanisms that allow us to retrieve the rich information provided by the motion of an agent. Both intrinsic and extrinsic motion may play an important role, however, in unravelling the mechanisms that mediate that information, it is wise to keep the two entirely different domains conceptually apart.

Within the domain of biological motion perception one important mechanism is the retrieval of articulated shape of the body. This process is – almost by definition – global as it is based on the relations between locations and motions of individual parts of the body. An interesting discussion in that context concerns the roles of global motion and global shape. Most of the earlier work emphasized the role of motion for the retrieval of shape (Hoffman & Flinchbaugh, 1982; Johansson, 1973; Webb & Aggarwal, 1982). The movements of small sets of dots that afford a partly rigid interpretation seem to benefit the retrieval of global shape. Cutting (Cutting, 1981) had shown already that perception of body shape from point-light displays is much easier when the dots are placed on the joints as compared to displaying off joint positions. In the latter case, pairs of dots do no longer retain constant distance in 3D and cannot be interpreted as rigid object parts anymore.

However, the argument can be reversed by putting the emphasis on the fact that observers are still able to perceive the articulation of a point-light walker to some degree – even when the dots are placed off joints. More recent work showed that the perception of a coherent shape that undergoes global motion is even possible when the stimulus is entirely deprived of any local motion whatsoever by assigning new dot positions on the body in each frame of the animation (Beintema & Lappe, 2002). The finding motivated a model that the incoming sensory data are tested against a prior template that implements knowledge about the articulation and overall shape of a human body (Lange & Lappe, 2006). The findings are interesting because they indicate that the mutual relations between global kinematic and global shape are not necessarily unidirectional. If local motion is not required to retrieve global shape, then the partial rigidity that can be obtained from spatial integration of local motions is also not

required. Note, however, that this of course does not mean that it is not used if available.

Yet another interesting question with respect to the relation between form and motion concerns the role of global motion after the articulated structure has been retrieved. Facing direction, actions, and probably even actor specifications such as gender or emotion can be retrieved from still stick figures (Coulson, 2004). Static stick figures can even give rise to the percept of motion, as has been demonstrated by countless cartoonists. In a photograph of a person it is in most cases easy to determine if the picture has been taken while the person was still or in motion (Kourtzi, Krekelberg, & van Wezel, 2008). So, once motion has served its role in filling in the missing sticks in a point-light animation, it is not clear to which degree it is still necessary and what kind of information it then conveys.

The discussion of the role of local motion for the retrieval of shape-independent information is entirely independent of the above dispute – which is all about how global coherence of shape and kinematic is obtained. Local motion by itself can contain interesting information, as discussed above, which is independent of global shape and global motion. Because of that independence it can serve as a generalizing invariant that encodes the presence of another living being, largely irrespective of its particular shape.

We introduced the term “life detector” for the invariant first suggested by Troje and Westhoff (2006) and further characterized by Chang (Chang & Troje, 2009a, 2009b) to express that it is independent of the particular shape of the animal in question. The term is not indisputable, though. The proposed visual filter is sensitive to ballistic properties of the feet which reflect an energetically efficient interaction with gravitational and inertial forces and therefore applies to legged, mid to large sized animals. However, it would not respond to fish, flying birds, snakes, and animals so small that their kinematics is dominated by air resistance rather than inertial forces. The invariant that we isolated from the foot motion of articulated animals is certainly not the only “life detection” mechanism. Others have yet to be identified – maybe in terms of “mid-level features” (Casile & Giese, 2005; Sigala, Serre, Poggio, & Giese, 2005) or “sprites”(Cavanagh, Labianca, & Thornton, 2001).

Global aspects of shape and motion, as well as local invariants like the one discussed above can contain semantic, socially relevant information about the category of the actor (human or animal? which kind?), and the action itself. The particular style of its execution can be exploited for further information about the identity of the actor and further details about the action. Again, both local and global information contribute in that respect. For instance, it has been shown that both global shape and global kinematics contribute to person identification and gender classification (Mather & Murdoch, 1994; Troje, 2002). Local aspects can convey emotional contents (e.g. Pollick, Paterson, Bruderlin, & Sanford, 2001) but global shape and motion can do so, too (e.g. Dittrich, et al., 1996).

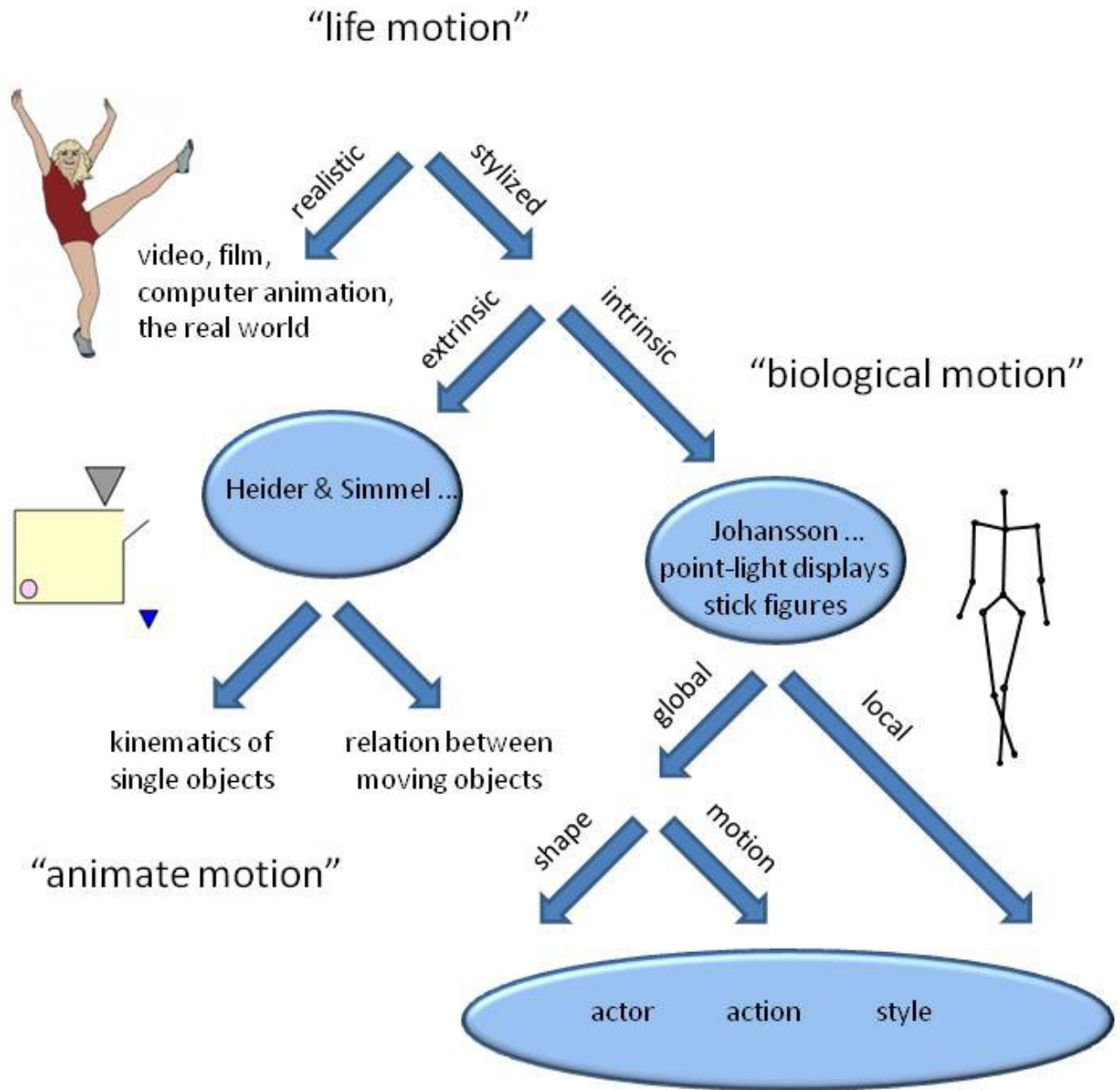
From this discussion it seems that information is encoded in biological motion in rather redundant ways. As with other stimuli, this can be exploited by the visual system to replace missing information in the stimulus and reconstruct attributes which are not directly exposed in it. If we see only the upper body of a person in motion, we can relatively accurately guess about what happens with the lower body. A 2D projection of a body in motion results in a strong hypothesis about its 3D structure. Temporal and spatial parameters of biological motion are related in specific ways and it has been shown that the visual system has knowledge about these relations and can exploit them to make predictions about missing parameters. For instance, the size of an animal and the cadence of its gait are related in a similar way as the length of a pendulum is related to the frequency with which it swings. When asked to estimate the size of a walking quadruped stick figure animal observers reproduce this relation, demonstrating that their visual system knows about it (Jokisch & Troje, 2003).

Little research has been done so far on the degree to which deviations from learned and expected correlations between kinematics and structure or between different movement parameters upsets the visual system. However, the reason why it is still not easy to generate artificial motion that fits a photorealistic avatar to a degree at which the visual system is seriously fooled into believing to see a real person is indicative for two things. On the one hand, it demonstrates the amazing abilities of the visual system to sense even the slightest inconsistencies within the sophisticated orchestra of the different aspects of shape and motion – maybe a trait that helped our hunting ancestors to identify weak or immature prey or to pick a healthy mate. On the other hand, it showcases the lack of understanding of the biomechanical details which determine these correlations – at least not to a degree at which we can come up with truly convincing generative models.

The non-linear relation between the realism achieved by a computer graphics characters and the comfort or discomfort that the observer experiences with it (sometimes termed the “uncanny valley”) might be a direct result of the level of consistency between the model that is applied to the stimulus and the stimulus itself. Models applied to aliens, monsters and cartoon figures probably represent much broader probability distributions as compared by the well-defined, peaked correlations that form the models that our visual system applies to real people. Biological motion point-light displays and stick figures evolved as a means to study life motion because they abstract the morphology of the figure to a point at which only few expectations can cause uncontrollable consistency problems. For this reason, biological motion point-light displays as introduced by Gunnar Johansson almost 40 years ago will be continued to be used in the future even though much more sophisticated depictions of human motion is available in today's age of computer animations and virtual realities.

Figure Caption

Fig. 1 depicts the suggested usage of the term “biological motion” and the context in which the field is positioned.



References

- Ahlstrom, V., Blake, R., & Ahlstrom, U. (1997). Perception of biological motion. *Perception*, *26*(12), 1539-1548.
- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends Cogn Sci*, *4*(7), 267-278.
- Atkinson, A. P., Dittrich, W. H., Gemmell, A. J., & Young, A. W. (2004). Emotion perception from dynamic and static body expressions in point-light and full-light displays. *Perception*, *33*, 717-746.
- Bardi, L., Regolin, L., & Simion, F. (2011). Biological motion preference in humans at birth: role of dynamic and configural properties. *Developmental Science*, *14*, 353–359.
- Beintema, J. A., & Lappe, M. (2002). Perception of biological motion without local image motion. *Proc Natl Acad Sci U S A*, *99*(8), 5661-5663.
- Bertenthal, B. I., & Pinto, J. (1994). Global processing of biological motions. *Psychological Science*, *5*(4), 221-225.
- Brand, M., & Hertzmann, A. (2000). *Style machines*. Paper presented at the Proceedings of SIGGRAPH 2000.
- Calvert, T. (2007). *Animating dance*. Paper presented at the Proceedings of Graphics Interface.
- Casile, A., & Giese, M. A. (2005). Critical features for the recognition of biological motion. *J Vis*, *5*(4), 348-360.
- Cavanagh, P., Labianca, A. T., & Thornton, I. M. (2001). Attention-based visual routines: sprites. *Cognition*, *80*(1-2), 47-60.
- Chang, D. H., & Troje, N. F. (2008). Perception of animacy and direction from local biological motion signals. *Journal of Vision*, *8*(5), 3 1-10.
- Chang, D. H., & Troje, N. F. (2009a). Acceleration carries the local inversion effect in biological motion perception. *J Vis*, *9*(1), 19 11-17. doi: 10.1167/9.1.19
/9/1/19/ [pii]
- Chang, D. H., & Troje, N. F. (2009b). Characterizing global and local mechanisms in biological motion perception. *J Vis*, *9*(5), 8 1-10. doi: 10.1167/9.5.8
/9/5/8/ [pii]
- Clarke, T. J., Bradshaw, M. F., Field, D. T., Hampson, S. E., & Rose, D. (2005). The perception of emotion from body movement in point-light displays of interpersonal dialogue. *Perception*, *34*(10), 1171.
- Coulson, M. (2004). Attributing emotion to static body postures: Recognition accuracy, confusions, and viewpoint dependence. *Journal of Nonverbal Behavior*, *28*(2), 117-139.
- Cutting, J. E. (1981). Coding theory adapted to gait perception. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(1), 71-87.
- 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.
- 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.
- Duncker, K. (1928). Über induzierte Bewegung. *Psychologische Forschung*, 12, 180-259.
- Farah, M. J. (1992). Is an object an object an object? Cognitive and neuropsychological investigations of domain specificity in visual object recognition. *Current Directions in Psychological Science*, 1(5), 164-169.
- Grossman, E. D., Donnelly, M., Price, R., Pickens, D., Morgan, V., Neighbor, G., & Blake, R. (2000). Brain areas involved in perception of biological motion. *J Cogn Neurosci*, 12(5), 711-720.
- Gurnsey, R., Roddy, G., Ouhana, M., & Troje, N. F. (2008). Stimulus magnification equates identification and discrimination of biological motion across the visual field. *Vision Research*, 48(28), 2827-2834. doi: S0042-6989(08)00466-5 [pii]
10.1016/j.visres.2008.09.016
- Gurnsey, R., Roddy, G., & Troje, N. F. (2010). Limits of peripheral direction discrimination of point-light walkers. *Journal of Vision*, 10(2). doi: 10.1167/10.2.15
- Heider, F., & Simmel, M. (1944). An experimental study of apparent behavior. *The American Journal of Psychology*, 57(2), 243-259.
- Hertzmann, A., O'Sullivan, C., & Perlin, K. (2009). *Realistic human body movement for emotional expressiveness*.
- Hirai, M., Chang, D. H. F., 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.
- Hirai, M., Saunders, D. R., & Troje, N. F. (2011). Allocation of attention to biological motion: local motion dominates global shape. *Journal of Vision*, 11(3), 1-11.
- Hiris, E., Humphrey, D., & Stout, A. (2005). Temporal properties in masking biological motion. *Percept Psychophys*, 67(3), 435-443.
- Hodgins, J. K., O'Brien, J. F., & Tumblin, J. (1998). Perception of human motion with different geometric models. *Visualization and Computer Graphics, IEEE Transactions on*, 4(4), 307-316.
- Hoffman, D. D., & Flinchbaugh, B. E. (1982). The interpretation of biological motion. *Biological Cybernetics*, 42(3), 195-204.
- Hohmann, T., Troje, N. F., Olmos, A., & Munzert, J. (2011). The influence of motor expertise and motor experience on action and actor recognition. *Journal of Cognitive Psychology*, 23(4), 403-415.
- Ikeda, H., Blake, R., & Watanabe, K. (2005). Eccentric perception of biological motion is unscalably poor. *Vision Res*, 45(15), 1935-1943.
- Jansson, G., Bergstroem, S. S., & Epstein, W. (1994). *Perceiving events and objects*.
- Jastorff, J., & Orban, G. A. (2009). Human functional magnetic resonance imaging reveals separation

and integration of shape and motion cues in biological motion processing. *Journal of Neuroscience* 29(22), 7315-7329. doi: 29/22/7315 [pii]

10.1523/JNEUROSCI.4870-08.2009

Jiang, Y., & He, S. (2007). Isolating the neural encoding of the local motion component in biological motion. *Journal of Vision*, 7(9), 551a.

Johansson, G. (1950). *Configurations in Event Perception*. Uppsala: Almqvist & Wiksell.

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

Johansson, G. (1974). Vector analysis in visual perception of rolling motion. *Psychological Research*, 36(4), 311-319.

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

Johnson, M. H. (2006). Biological motion: a perceptual life detector? *Curr Biol*, 16(10), R376-377.

Jokisch, D., & Troje, N. F. (2003). Biological motion as a cue for the perception of size. *Journal of Vision*, 3(4), 252-264.

Knoblich, G., & Flach, R. (2001). Predicting the effects of actions: interactions of perception and action. *Psychological Science*, 12(6), 467-472.

Kourtzi, Z., Krekelberg, B., & van Wezel, R. J. A. (2008). Linking form and motion in the primate brain. *Trends in Cognitive Sciences*, 12(6), 230-236.

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.

Kriegstein, K. V., & Giraud, A. L. (2004). Distinct functional substrates along the right superior temporal sulcus for the processing of voices. *Neuroimage*, 22(2), 948-955.

Lange, J., & Lappe, M. (2006). A model of biological motion perception from configural form cues. *Journal of Neuroscience* 26(11), 2894-2906. doi: 26/11/2894 [pii]

10.1523/JNEUROSCI.4915-05.2006

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.

Mather, G., Radford, K., & West, S. (1992). Low-level visual processing of biological motion. *Proc R Soc Lond B Biol Sci*, 249(1325), 149-155.

Montepare, J. M., & Zebrowitz-McArthur, L. (1988). Impressions of people created by age-related qualities of their gaits. *Journal of Personality and Social Psychology*, 55(4), 547.

Munzert, J., Hohmann, T., & Hossner, E. J. (2010). Discriminating throwing distances from point-light displays with masked ball flight. *European Journal of Cognitive Psychology*, 22(2), 247-264.

Pilz, K. S., Bennett, P. J., & Sekuler, A. B. (2010). Effects of aging on biological motion discrimination. *Vision Research*, 50(2), 211-219.

- Pollick, F. E., Paterson, H. M., Bruderlin, A., & Sanford, A. J. (2001). Perceiving affect from arm movement. *Cognition*, *82*(2), B51-61.
- Puce, A., & Perrett, D. I. (2003). Electrophysiology and brain imaging of biological motion. *Philos Trans R Soc Lond B Biol Sci*(358), 435-445.
- Runeson, S., & Frykholm, G. (1981). Visual perception of lifted weight. *Journal of Experimental Psychology: Human Perception and Performance*, *7*(4), 733-740.
- Runeson, S., & Frykholm, G. (1983). Kinematic specification of dynamics as an informational basis for person-and-action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General*, *112*(4), 585-615.
- Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences*, *4*(8), 299-309.
- Sigala, R., Serre, T., Poggio, T., & Giese, M. (2005). Learning features of intermediate complexity for the recognition of biological motion. *Artificial Neural Networks: Biological Inspirations—ICANN 2005*, 241-246.
- Simion, F., Regolin, L., & Bulf, H. (2008). A predisposition for biological motion in the newborn baby. *Proc Natl Acad Sci U S A*, *105*(2), 809-813.
- Thompson, B., Hansen, B. C., Hess, R. F., & Troje, N. F. (2007). Peripheral vision; good for biological motion, bad for signal noise segregation? *Journal of Vision*, *7*(10), 1-7.
- Thompson, B., Troje, N. F., Hansen, B. C., & Hess, R. F. (2008). Amblyopic perception of biological motion. *J Vis*, *8*(4), 22 21-14. doi: 10.1167/8.4.22
- /8/4/22/ [pii]
- Troje, N. F. (2002). Decomposing biological motion: a framework for analysis and synthesis of human gait patterns. *Journal of Vision*, *2*(5), 371-387.
- Troje, N. F. (2008). Retrieving information from human movement patterns. In T. F. Shipley & J. M. Zacks (Eds.), *Understanding Events: How Humans See, Represent, and Act on Events* (pp. 308-334). Oxford: Oxford University Press.
- Troje, N. F., & Chang, D. H. F. (in press). Shape-independent processes in biological motion perception. [Book Chapters]. In: *Shiffrar, M. and Johnson, K. (eds.) Visual Perception of the Human Body in Motion: Findings, Theory, and Practice. Oxford University Press.*
- Troje, N. F., & Westhoff, C. (2006). The inversion effect in biological motion perception: evidence for a "life detector"? *Curr Biol*, *16*(8), 821-824.
- Troje, N. F., Westhoff, C., & Lavrov, M. (2005). Person identification from biological motion: effects of structural and kinematic cues. *Percept Psychophys*, *67*(4), 667-675.
- Ullman, S. (1979). The interpretation of structure from motion. *Proceedings of the Royal Society of London. Series B. Biological Sciences*, *203*(1153), 405.
- Unuma, M., & Takeuchi, R. (1993). Generation of human motion with emotion. *Computer Animation '93*, 77-88.
- Vaina, L. M., Solomon, J., Chowdhury, S., Sinha, P., & Belliveau, J. W. (2001). Functional neuroanatomy of biological motion perception in humans. *Proc Natl Acad Sci U S A*, *98*(20), 11656-11661.

Vallortigara, G., & Regolin, L. (2006). Gravity bias in the interpretation of biological motion by inexperienced chicks. *Curr Biol*, *16*(8), R279-280.

Wang, L., & Jiang, Y. (in press). Life motion signals lengthen perceived temporal duration. *Proceedings of the National Academy of Sciences*.

Wang, L., Zhang, K., He, S., & Jiang, Y. (2010). Searching for Life Motion Signals. *Psychological Science*, *21*(8), 1083.

Webb, J. A., & Aggarwal, J. K. (1982). Structure from motion of rigid and jointed objects. *Artificial Intelligence*, *19*(1), 107-130.

Westhoff, C., & Troje, N. F. (2007). Kinematic cues for person identification from biological motion. *Percept Psychophys*, *69*(2), 241-253.