

Perception of biological motion as motion-from-form

Introduction

In the second half of the 19th century, soon after the invention of photography, photographers and scientists realized that it is possible to capture the movements of humans or animals in a sequence of still-frame pictures. The American photographer Eadward Muybridge pioneered this approach. In 1878 he became world-famous for his series of photographs of a galloping horse. These images showed for the first time that there is a moment in the horse's gait in which all four hoofs are in the air at once (■ Fig. 1). In the following years, the French physiologist Etienne-Jules Marey adapted this technique to study human movement. In order to capture purely the motions of the limbs, he clothed actors in black suits with white bands attached to the arms and legs and photographed them against a black background (■ Fig. 1, lower left panel). The sequence of still frames then gave him enough information to trace and analyze the movements of the body. About 100 years later, Gunnar Johansson [13] in Sweden discovered that such abstracted information is actually sufficient as a visual stimulus for perception of the movement of the actor when it is synthesized into a movie.

The stimulus that Johansson used subsequently came to be called the point-light display. It consisted of the motion of a small number of light points attached to the major joints of the body of an actor (■ Fig. 1, lower right panel). Johansson demonstrated that human observers can

perceive highly complex features of human movement and action from this very impoverished visual stimulus. He called the ability to perceive the actor and its actions the perception of biological motion. Later studies showed that the movement of animals can also be perceived from point-light displays [19] and that observers are even able to recognize individuals or the gender of a person [7, 14].

This ability appeared astounding to many since the stimulus seemed so impoverished and devoid of almost all visual information about the actor. In Johansson's studies, immediate perception of biological motion occurred only when the point-light stimulus was set in motion. A single image of a point-light figure was insufficient to elicit the perception of a human figure. Thus Johansson concluded that the information in a point-light display is carried mainly by the motion of the points over time. Since then, biological motion perception has often been regarded as a highly specialized form of motion analysis, i.e. a perception of form-from-motion. However, research on biological motion perception over the last 10 years or so has provided evidence for a rather different view, namely that biological motion is derived from the analysis of sequences of body postures. In this view, biological motion is motion-from-form processing. Psychophysical, physiological and computational studies support this view.

Form and motion in point-light displays

In point-light displays, a small number of light points are shown in a movie or computer animation (■ Fig. 2a). These light points represent the position and movement of the major joints of the human body. Johansson's original displays were constructed by filming actors who had small light bulbs attached to their bodies. Later studies have sometimes used a computer program that simulates the joint movements of a walking human figure instead [6].

Point-light displays contain information about the position and the motion vectors of the joints (■ Fig. 2b). The motion information is directly specified by the change of position (apparent motion) of the light points over frames. Information about the form of the body, on the other hand, is largely removed because the outline of the body is not visible. Some limited form information is retained, however, in the positioning of the light points on the joints. In principle, a static image of a single frame from a point-light animation could provide enough information to estimate the body posture, if one knows how to connect the correct points with lines.

The percept generated by point-light biological motion encompasses both the form of a human figure and the motion of its limbs. It therefore involves both form and motion recognition. However, there are two routes by which the visual system might arrive at this percept. The first

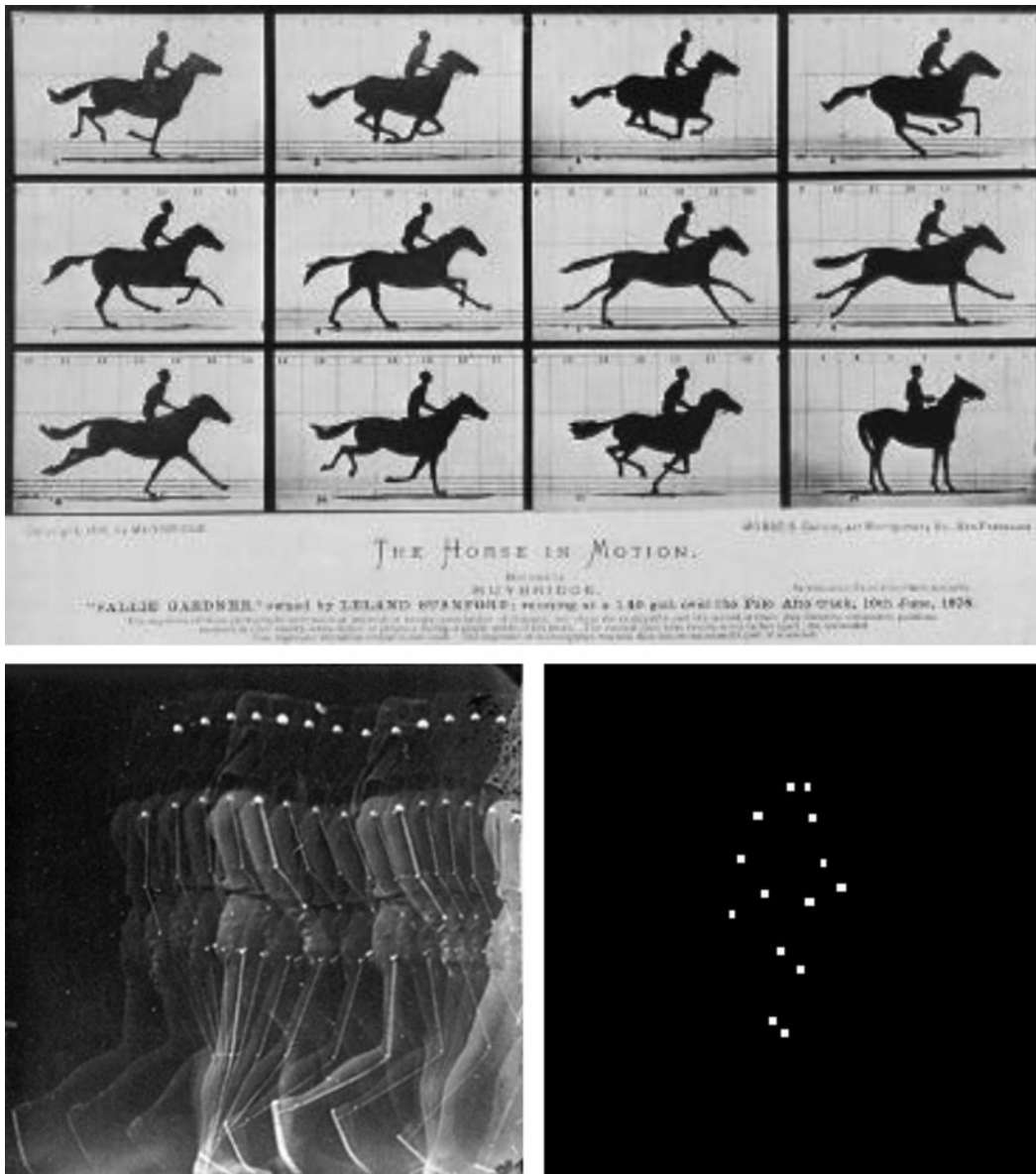


Fig. 1 ◀ *Top* A sequence of static images of a galloping horse shot by Eadweard Muybridge in the late 19th century on a racetrack with a series of cameras taking pictures one after another. The images show in detail the stages of movement. *Bottom left* Image taken by the French physiologist Etienne-Jules Marey of an actor wearing a dark suit with reflecting stripes and dots. In this superimposed sequential exposure on a single image plate the flow of movements of the body can be visualized and analyzed. *Bottom right* A single frame from a typical point light biological motion stimulus. An actress walks towards the camera. Only the point lights at the head, shoulders, elbows, hands, hips, knees and ankles are visible

starts by computing motion vectors from the light points (right panel of **Fig. 2b**). The pattern of motion vectors is then analyzed and interpreted, perhaps in conjunction with knowledge or expectations about the form and movement of a human body. This motion-based approach considers biological motion perception a variant of form-from-motion perception.

The alternative route to biological motion perception assumes that the temporal evolution of the static form information provided by each single frame image may be used over time to integrate form cues across views. This approach starts by computing human form information from the positions of the light points in a static im-

age (left panel of **Fig. 2b**). It then calculates the motion of body parts from changes in the form. In this form-based approach, biological motion perception is the recognition of dynamic form. I will call this motion-from-form perception.

In the motion-from-form approach, the visual input is first used to estimate the form and posture of the body. Motion is then derived from the changing body form. The motion-from-form approach does not require local motion vectors. Instead, it attempts to find positions of points on the body rather than the motion vectors of these points. This raises the question of whether point light displays contain enough information to sup-

port human form recognition without using motion vectors, and how this process might be implemented in the brain.

Point-light displays without motion

Is there enough information in the position of the points of point-light displays to support form analysis? If so, why did observers not recognize a moving figure from a static Johansson display? The first answer to these questions is that later research has shown that whether or not a static point light display is recognized as a human figure depends on the posture that is displayed [8, 26]. Postures in which

the extremities are extended are more easily recognized than those in which the extremities are close to the body. Secondly, there are at least two possibilities as to why setting the display in motion gives more information than a single image. First, in a sequence of images form information provided by each single image can be accumulated over time. This means that a number of body postures are displayed over time and each provides the system with more constraints on the interpretation of the image series. Because each single image carries very little form information, such a temporal accumulation might be an essential requirement to see the walking figure. Secondly, even when position information is the primary cue that is used to recognize the figure, a sequence of images allows the observer to also estimate the action of the figure. Recognizing the action may be a fundamental part of the spontaneous recognition of biological motion in point-light displays.

The above argument illustrates a problem in investigating the contributions of motion and form to biological motion recognition. Because the point-light display contains both form and motion it is difficult to estimate the respective role of each. Beintema and Lappe [2] used a limited lifetime technique to create point-light stimuli in which the use of image motion information is reduced (■ Fig. 2c). These stimuli directly pitted motion and form information against each other. A small number of light points were placed on the outline of the body rather than on the joints. Each light point remained at its position on the body for only a limited time. Thereafter, the point was extinguished and a new one was created at a different position. In these stimuli, the form of the body is sampled over time more completely than in classic point-light stimuli. Each individual image, however, gives only very limited form information. The amount of form information can be adjusted by varying the number of dots displayed simultaneously. The amount of image motion information, on the other hand, can be adjusted by varying the lifetime of each dot. If a dot stays on the same position on the limb for two or more frames this dot generates an apparent motion signal. If the lifetime is restricted to

only a single frame, no individual point creates apparent motion in the direction of the limb movement.

With these stimuli, biological motion recognition was possible even with a lifetime of one frame, i.e. in the absence of apparent motion signals of the limb movement [2]. This suggests that form information, although limited in any single image, can be exploited in a sequence of images. When point lifetime was increased, image motion signals were added at the cost of a slower sampling of body positions. In this case, performance in direction or coherence discrimination dropped. This suggests that position and form cues are more important in these tasks than image motion cues. Finally, these stimuli allow us to look at the role of the temporal integration of position information. We have argued above that a richer sample of position signals can be obtained from a sequence of images than from any single image alone. This temporal integration might be associated with the greater ability to spontaneously recognize a point-light display in motion than a still frame. However, in that case there are also motion signals that may contribute to the percept. In the stimuli of Beintema and Lappe [2], a richer sample of position signals may be generated in an image sequence without setting the figure in motion. This occurs when a single static posture of a human body is displayed with limited lifetime temporal sampling. In this condition, naive subjects were able to recognize the figure purely from position signals presented over time. This suggests that the temporal integration of position signals may be a viable mechanism for figure recognition and for a subsequent recognition of biological motion.

The observation that the human form can be recognized from point-light displays even when the figure remains static underlines the importance of the form information. However, this finding raises the question of how the motion of the body is recognized, and how the recognition of body motion can be investigated. Many psychophysical studies on biological motion perception have used direction discrimination to investigate perception. Subjects had to discriminate whether a walker was facing to the left or to the right. However, since this discrimination

e-Neuroforum 2012 · 3:67–73
DOI 10.1007/s13295-012-0032-y
© Springer-Verlag 2012

M. Lappe

Perception of biological motion as motion-from-form

Abstract

The recognition of the movements and actions of others is of great importance for social interaction. The visual motion pattern projected on the retina when watching somebody else act is called biological motion. Because of the many degrees of freedom of the body, biological motion is a relatively complicated motion pattern, much more variable, for example, than optic flow or object motion. The regularities of biological motion are contained in its relationship to the body, i.e. in the constraints imposed by the articulation of the limbs on the movement of the body parts. The neural mechanisms of biological motion perception, therefore, take body form information into account. I describe a model of biological motion perception that starts from a representation of body form and posture and retrieves biological motion as the transformation of the body posture over time. Essentially, this proposes a ventral pathway to motion perception that is distinct from the other motion pathways in the dorsal stream, and specialized for body motion.

Keywords

Visual perception · Biological motion · Point-light walker · Action recognition · Form-from-motion

can be performed on a single static posture it does not truly test motion processing. Beintema and Lappe [3] suggested a discrimination between forward and backward walking instead, and compared performance in both discrimination tasks (and a third based on the structural coherence of the body) for a variety of presentation durations, point numbers and point lifetimes. They found that discrimination performance was mainly determined by the total number of points seen during stimulus presentation, and less by how many points were in any given frame, or how long each frame lasted, or whether the points moved consistently from frame to frame with lifetimes greater than one (■ Fig. 3). Moreover, the discrimination between forward and backward walking (body motion discrimination) required

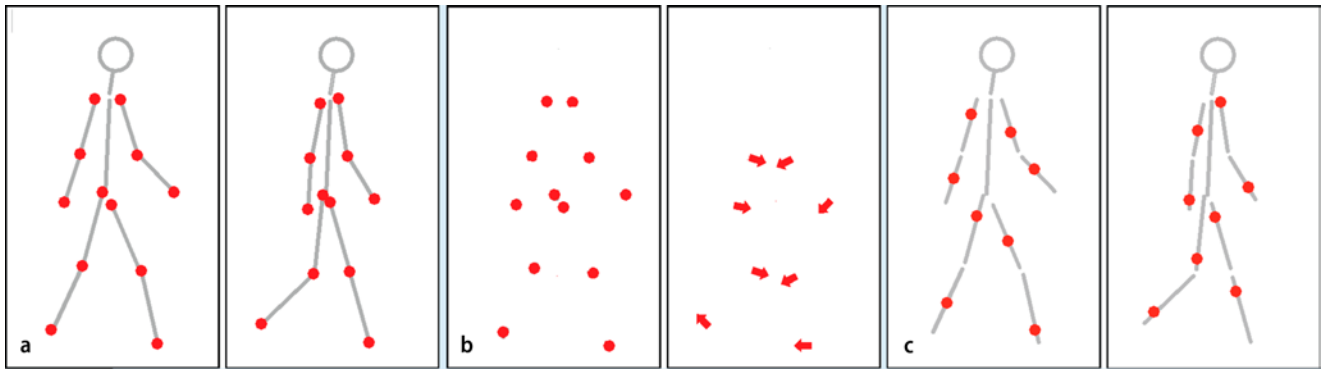


Fig. 2 **a** Two frames of a point-light walker stimulus. The stimulus consists of 12 points attached to the major joints of the body. The grey lines are not part of the stimulus but serve here to illustrate the body structure. From one frame to the next the body posture changes in accordance with a forward step. **b** Two types of information in this stimulus: position (*left*) and motion (*right*) of the points. The position information is available in each single frame. The motion vectors result from the apparent motion of a point between subsequent frames. **c** Two frames of a point-light walker devoid of the motion signals depicted in **b** [2]. The points are distributed over the limbs in each frame but the distribution is randomized from one frame to the next. Thus, there is no basis for apparent motion signals in the direction of the limb movement

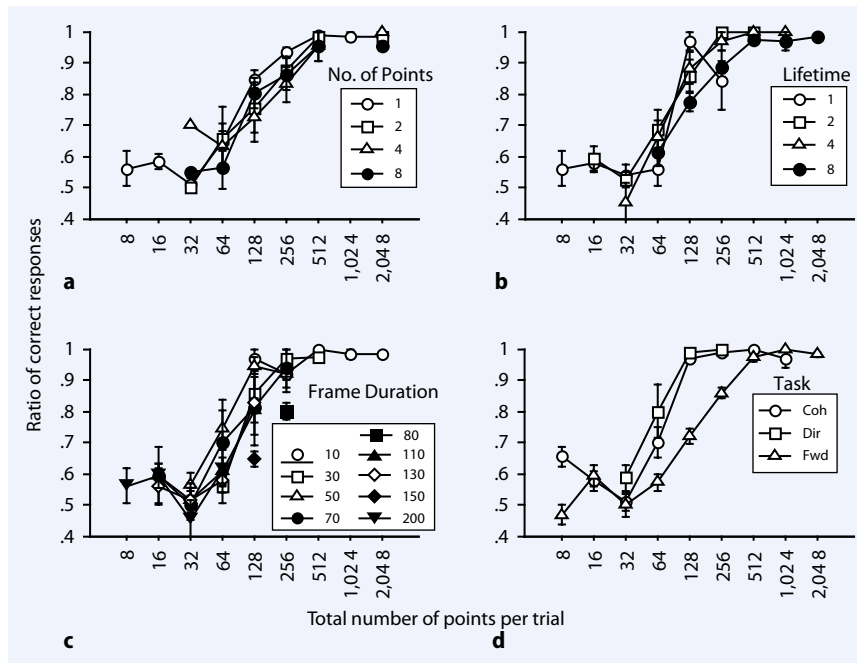


Fig. 3 **a** The figure combines data from several experiments with the limited lifetime stimulus (Fig. 2c) of Beintema et al. [2, 3]. In these experiments observers had to recognize either the structural coherence of the body, the facing direction (left or right), or the motion direction (forward or backward walking) for various combinations of the stimulus parameters (number of points per frame, lifetime of each point and presentation duration of each frame). Each of the four panels shows the same data, but split into different curves depending on number of points, frame duration, lifetime and task. The x-axis in each panel gives the total number of points in each trial, i.e. the number of points per frame times the number of frames in the trial. The figure shows that performance is mostly determined by the total number of points per trial. For example, performance saturates at around 256 points per trial, irrespective of whether this number is reached by one point per frame and 256 frames at 10-ms frame duration, or by eight points per frame for 32 frames at 50-ms frame duration. This is consistent with the idea that facing direction and coherence rely on body posture, to which each point contributes the same amount of information. Only the split by task (*lower right panel*) shows a clear difference. The discrimination between forward and backward walking (motion) needs about twice as many points as the discrimination of facing and coherence. This is consistent with the idea that motion is determined from the analysis of body posture in (at least two) successive frames

about twice the number of points than discrimination based on body form (facing and coherence discrimination). These findings are consistent with the idea that body motion is derived from body form processing and needs at least two sequential postures.

Biological motion from sequential posture analysis

Beintema and Lappe [2] proposed that biological motion perception may be performed by an analysis of sequential posture information, obtained from position signals of points on the body. This could be done via dynamic form templates that accumulate the evidence for human form over time, while allowing for a dynamic change in the form of the body. Lange and Lappe [16] transformed this idea into a biologically plausible neurocomputational model (Fig. 4) that captures many of the psychophysical and physiological properties of biological motion perception [15, 16, 17, 18]. This model starts with a set of template cells that each represent a particular posture of the human body. Their activities are determined from the match of each single frame of the stimulus to the preferred posture of each neuron. As the stimulus moves and the body posture changes, a sequence of body posture cells is activated one after another. The estimation of the walking movement is then performed by neurons that respond specifically to one (forward) or

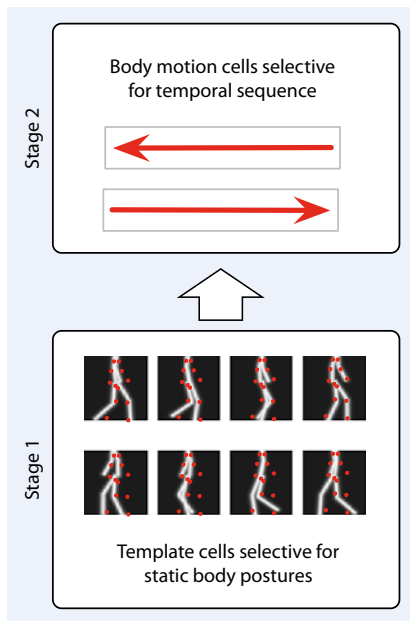


Fig. 4 ▲ The model of biological motion perception from sequential posture analysis [16] consists of two processing stages. The first stage contains neurons selective for the posture of the body seen from a particular view point. Each cell can be considered as a template for the body view in this posture. The cell responds according to the match of the stimulus dots with its template on a frame-by-frame basis. The discrimination of the facing direction (left or right) of the stimulus is computed by comparing the maximum template responses for each direction. The second stage computes the temporal sequence of activation of the template cells to generate selectivity for body motion. The discrimination between forward and backward walking is computed by comparing the activations for the two possible motion sequences. The selectivity to posture in the template cells relates to body form processing areas of the brain (extrastriate and fusiform body areas), and the selectivity to body motion relates to biological motion selectivity in the superior temporal sulcus

the other (backward) sequence of activities. The discrimination of the facing direction of the stimulus (walking leftward or rightward), on the other hand, is performed directly on the body posture templates, by finding those templates that are most active for a given stimulus. Hence, the model proposes a two-stage recognition scheme, in which first the posture of the body and then the motion of the body are analyzed. This scheme predicts a neural representation of body posture in the brain and a neural representation of body motion.

Body posture representations in the brain

In the human brain, two areas have been identified that respond selectively to visual images of the human body: the extrastriate body area (EBA) [9] and the fusiform body area (FBA) [24]. Selectivity for body form and posture has also recently been found in monkey temporal cortex [34]. These posture representations may form the basis for the recognition of biological motion. Indeed, in human fMRI studies with point-light walkers, the body-selective areas are activated also by point-light stimuli that convey only very limited information about body structure [4, 11, 33]. Moreover, using the limited lifetime technique described above, Michels et al. [21] found activations in body form-selective areas also for static postures of point-light stimuli. In the monkey, Vangeneugden et al. [34] showed that most temporal cortex neurons that responded to sequences of stick figures of a human body in walking motion actually responded to particular static postures within this sequence.

If biological motion perception is based on such posture-selective neurons in a two-stage process, then the posture representation should contain information about the facing direction of the body. Indeed, Vangeneugden et al. [34] found cells specific for particular facing directions and showed that a support vector machine classification analysis based on the temporal cortical population responses was very effective in discriminating facing direction. In humans, Michels et al. [22] provided evidence that different facing directions of point-light stimuli are represented in distinct patches in the fusiform gyrus.

Another way to show neural specificity to particular stimuli is via the aftereffect method. In an aftereffect experiment, a stimulus is shown for a long duration during which cells selective for the properties of this stimulus are fatigued. When a neutral stimulus is shown immediately afterwards the percept is often one of the opposite of the previously presented stimulus. This is taken as evidence that the original stimulus is encoded in a dedicated population of neurons. Theusner et

al. [29] performed such an aftereffect experiment using a walker facing in one direction as adaptor and a superposition of two walkers facing in two directions as a neutral stimulus. The results showed that facing direction can be selectively adapted, confirming that the neural representation of walking contains facing-specific populations. Other aftereffect studies have shown further properties of point-light walkers that are coded in specific representations including gender [31] and heading direction [12] of the walker. Also walking direction (forward vs. backward walking) shows aftereffects such that, for example, prolonged viewing of a forward walking walker induces the percept of backward walking in subsequently shown ambiguous [29] or static [1] stimuli. This is particularly important for the mechanisms of biological motion perception from posture sequence analysis because the difference between forward and backward walking lies in the temporal order of the posture sequence, and thus allows us to investigate the second stage of the above model, the body motion level.

Body motion representations in the brain

Neuroimaging studies in humans have shown selectivity to biological motion in the superior temporal sulcus (STS) [4, 11, 33]. This is consistent with early studies in monkeys that showed selectivity to body motion and point-light walkers in the superior temporal polysensory area [23]. Besides STS, activation by biological motion stimuli has also been reported in the above-mentioned body areas, in premotor cortex, motion areas hMT+ and KO and in the cerebellum [25, 28, 33]. The STS has reciprocal connections to areas of the form pathway in ventral cortex and of the motion pathway in dorsal cortex. Input from the ventral body posture representations could, therefore, be used in the STS to analyze the temporal order of the posture sequence and determine body motion. Indeed, activation of human STS was found not only for classical point-light stimuli but also for point-light walkers devoid of local motion signals, for which body motion is available only from posture sequence analysis [21, 22]. Con-

versely, in a study that manipulated body form information by separating the limbs from one another while keeping their motion intact, activation in the STS was reduced, confirming that body form information was important for driving STS activation [30].

Different representations for body form and body motion were also identified in the monkey [34]. A subset of temporal cortex neurons responded to the sequence of body motion postures (i.e. body motion) stronger than to individual postures. These neurons were found predominantly in the upper bank of the STS, whereas the posture-selective neurons were more frequent in the lower bank.

Motion-from-form

The model presented here assumes a representation of body posture from neurons selective for body forms. Body motion then induces a temporal variation of activity in this “posture space”. Biological motion perception can then be performed by applying motion detection mechanisms to this “posturo-temporal” signal. The result is a biological motion detector that is based upon body form transformation, i.e. a motion-from-form pathway. This model is supported by many experimental findings from psychophysical studies, observations of aftereffects, neuroimaging studies and electrophysiological experiments in monkeys. In addition, there are reports of patients with deficits in general motion perception that can nonetheless recognize biological motion [20, 32]. All of this suggests that biological motion perception can progress via a first analysis of the form of the body and a subsequent analysis of the motion of the body from the change of the body form or posture over time. This constitutes a route to body motion perception that does not involve the regular motion pathway of the brain but rather a motion mechanism acting on top of form analysis. Whereas regular motion perception is based on the variation of the spatial distribution of luminance over time, the motion-from-form pathway to biological motion is based on the variation of posture over time.

This does not preclude, however, that the regular motion pathway also contrib-

utes to biological motion perception. For example, it may be that motion signals from the individual points of a point-light walker are combined into a complex motion pattern that signals biological motion (e.g. [10, 13]). Also, the motion trajectories of individual point lights, such as the feet, can convey particular aspects of biological motion perception, for example the facing direction [27], and support a general percept of animacy [5]. Biological motion may be too complex and multi-faceted a percept to be explained by any simple single mechanism. However, for the current understanding of the motion processing pathways in the brain it demonstrates a route to motion that bypasses regular luminance-based motion detection and instead works through the analysis of form changes signaled via body posture representations in the brain's form pathway.

Corresponding address

Prof. Dr. M. Lappe

Institute of Psychology & Otto Creutzfeldt Center for Cognitive and Behavioral Neuroscience, University of Muenster
Fliednerstr. 21, 48149 Muenster
Germany
mlappe@uni-muenster.de

Prof. Dr. M. Lappe Markus Lappe studied physics in Marburg and Tübingen, where he received his PhD based on research in computational neuroscience of optic flow in the lab of Josef Raucherer at the National Institutes of Health, USA (1990–1992). In 1993 he joined the lab of Klaus-Peter Hoffmann at the Ruhr-University Bochum where he worked on computational and experimental neuroscience of visual motion and eye movements. In 1999 he received the Bio-Future Prize of the German Ministry for Science and Technology. In addition to being a founding member of the Otto Creutzfeldt Center for Cognitive and Behavioral Neuroscience, he has been a full professor of experimental psychology at the University of Muenster since 2001.

Conflict of interest. No statement made.

References

1. Barraclough N, Jellema T (2011) Visual aftereffects for walking actions reveal underlying neural mechanisms for action recognition. *Psychol Sci* 22(1):87–94
2. Beintema JA, Lappe M (2002) Perception of biological motion without local image motion. *Proc Natl Acad Sci U S A* 99(8):5661–5663
3. Beintema JA, Georg K, Lappe M (2006) Perception of biological motion from limited lifetime stimuli. *Percept Psychophys* 68(4):613–624

4. Beauchamp MS, Lee KE, Haxby JV, Martin A (2003) fMRI responses to video and point-light displays of moving humans and manipulable objects. *J Cogn Neurosci* 15:991–1007
5. Chang DHF, Troje NF (2007) Perception of animacy and direction from local biological motion signals. *J Vis* 8(5):3.1–10
6. Cutting JE (1978) A program to generate synthetic walkers as dynamic point-light displays. *Behav Res Meth Instrum Comput* 10(1):91–94
7. Cutting JE, Kozlowski LT (1977) Recognizing friends by their walk: Gait perception without familiarity cues. *Bull Psychonom Soc* 9:353–356
8. Cutting JE, Moore C, Morrison R (1988) Masking the motions of human gait. *Percept Psychophys* 44(4):339–347
9. Downing PE, Jiang Y, Shuman M, Kanwisher N (2001) A cortical area selective for visual processing of the human body. *Science* 293(5539):2470–2473
10. Giese MA, Poggio T (2003) Neural mechanisms for the recognition of biological movements. *Nat Rev Neurosci* 4(3):179–192
11. Grossman ED, Blake R (2002) Brain areas active during visual perception of biological motion. *Neuron* 35(6):1167–1175
12. Jackson S, Blake R (2010) Neural integration of information specifying human structure from form, motion and depth. *J Neurosci* 30(3):838–848
13. Johansson G (1973) Visual perception of biological motion and a model for its analysis. *Percept Psychophys* 14:201–211
14. Kozlowski LT, Cutting JE (1977) Recognizing the sex of a walker from a dynamic point-light display. *Percept Psychophys* 21(6):575–580
15. Lange J, Georg K, Lappe M (2006) Visual perception of biological motion by form: a template-matching analysis. *J Vis* 6(8):836–849
16. Lange J, Lappe M (2006) A model of biological motion perception from configurational form cues. *J Neurosci* 26(11):2894–2906
17. Lange J, Lappe M (2007) The role of spatial and temporal information in biological motion perception. *Adv Cogn Psychol* 3(4):419–429
18. Lange J, Lappe M (2010) Dynamic form templates determine sensitivity to biological motion. In: Wang R, Gu F (ed) *Advances in cognitive neurodynamics*, Vol 2. Springer, pp 409–414
19. Mather G, West S (1993) Recognition of animal locomotion from dynamic point-light displays. *Perception* 22(7):759–766
20. McLeod P, Dittrich W, Driver J et al (1996) Preserved and impaired detection of structure from motion by a ‘motion-blind’ patient. *Vis Cogn* 3(4):363–391
21. Michels L, Lappe M, Vaina LM (2005) Visual areas involved in the perception of human movement from dynamic form analysis. *NeuroReport* 16(10):1037–1041
22. Michels L, Kleiser R, de Lussanet MHE et al (2009) Brain activity for peripheral biological motion in the posterior superior temporal gyrus and the fusiform gyrus: Dependence on visual hemifield and view orientation. *Neuroimage* 45(1):151–159
23. Oram MW, Perrett DI (1994) Responses of anterior superior temporal polysensory (stpa) neurons to biological motion stimuli. *J Cogn Neurosci* 6(2):99–116
24. Peelen MV, Downing PE (2005) Selectivity for the human body in the fusiform gyrus. *J Neurophysiol* 93(1):603–608
25. Peuskens H, Vanrie J, Verfaillie K, Orban GA (2005) Specificity of regions processing biological motion. *Eur J Neurosci* 21(10):2864–2875

-
26. Reid R, Brooks A, Blair D, Zwan R van der (2009) Snap! recognising implicit actions in static point-light displays. *Perception* 38(4):613–616
 27. Saunders DR, Suchan J, Troje NF (2009) Off on the wrong foot: local features in biological motion. *Perception* 38(4):522–532
 28. Saygin AP, Wilson SM, Hagler DJ et al (2004) Point-light biological motion perception activates human premotor cortex. *J Neurosci* 24(27):6181–6188
 29. Theusner S, de Lussanet MHE, Lappe M (2011) Adaptation to biological motion leads to a motion and a form aftereffect. *Atten Percept Psychophys* 73(6):1843–1855
 30. Thompson JC, Clarke M, Stewart T, Puce A (2005) Configural processing of biological motion in human superior temporal sulcus. *J Neurosci* 25(39):9059–9066
 31. Troje NF, Sadr J, Geyer H, Nakayama K (2006) Adaptation aftereffects in the perception of gender from biological motion. *J Vis* 6(1534–7362 (Electronic)):850–857
 32. Vaina LM, Lemay M, Bienfang DC et al (1990) Intact biological motion and structure from motion perception in a patient with impaired motion mechanisms: a case study. *Vis Neurosci* 5:353–369
 33. Vaina LM, Solomon J, Chowdhury S et al (2001) Functional neuroanatomy of biological motion perception in humans. *Proc Natl Acad Sci U S A* 98(20):11656–11661
 34. Vangeneugden J, De Maziere PA, Van Hulle MM et al (2011) Distinct mechanisms for coding of visual actions in macaque temporal cortex. *J Neurosci* 31(2):385–401