

Ideal-observer-model and psychophysical experiments on the role of form information in biological motion perception

Joachim Lange, Karsten Georg, Markus Lappe

Department of Psychology, Westf. Wilhelms-University Münster, Germany

Abstract. 'Biological motion perception' refers to the impressive ability of human observers to visually identify the motion of humans or animals solely from the moving patterns of a small number of light points attached to the body. Although the first experiments concerning the perception of biological motion already took place in 1973 [1] the perceptual mechanisms are still poorly understood. Based on experiments with a novel biological motion stimulus Beintema and Lappe [2] recently proposed that the perception of biological motion relies more on form than on motion signals. We developed an ideal-observer-model which is based on form information only. In various forced-choice experiments we compared the model's performance with that of human observers in psychophysical studies. The model results showed striking similarities with the data from human subjects. These findings lend additional support to the idea that biological motion perception is based on an analysis of sequential poses each derived from form signals.

1 Introduction

A walking human person produces a highly complex visual motion pattern. However, despite its non-rigidness and its many degrees of freedom this pattern can be recognized by human observers in a fraction of a second. Johansson [1] revealed that this is even true when the visible information is reduced to only a few light points fixed on the joints of the walker. The information transmitted by this 'point-light' display, which is commonly presented as a computer animation [3], can be subdivided into motion and position signals (figure 1a). A single frame of this animation provides form information via the joint positions. A sequence of frames provides motion information via apparent motion signals of the individual points. Since a single frame does not induce the percept of biological motion in naive observers, many studies and models argued that the rapid recognition of biological motion is based on motion signals [1, 4]. Interestingly, however, some patients with lesions in the motion processing areas of the brain are impaired in perception of general aspects of image motion but not in the recognition of biological motion [5, 6].

Beintema and Lappe hypothesized upon these findings that the recognition of biological motion is based on spatiotemporal integration of form information

rather than directly on motion signals [2]. They created a new biological motion stimulus by placing light points at random positions on the extremities rather than on the joints, and then removed local motion signals by jumping points randomly to new positions on the body for each animation frame. Psychophysical

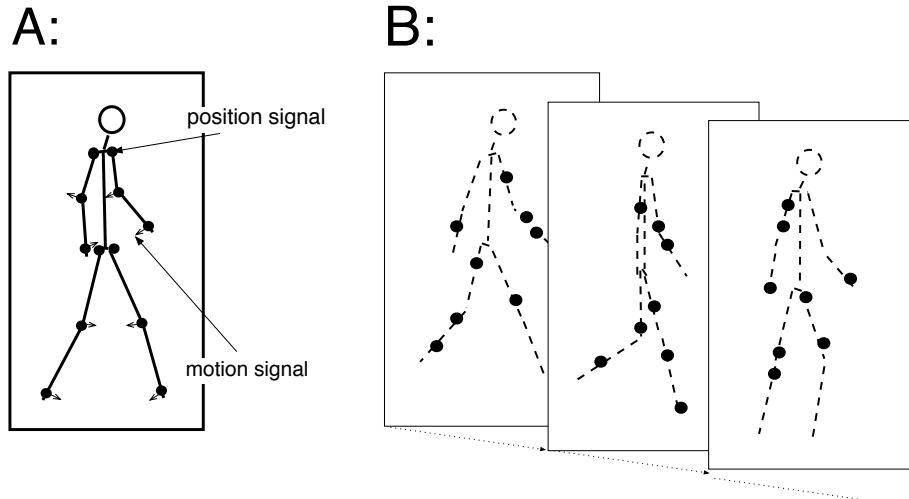


Fig. 1. A: Subdivision of the signals from the walker into position and motion components. B: The single-frame-lifetime (SFL) stimulus consisted of dots that changed their position on the limbs randomly from frame to frame

studies with these 'single-frame-lifetime' (SFL) stimuli showed that biological motion was still perceived from this stimulus, and that two classical 2AFC tasks, direction (SFL-Walker walking either to the right or to the left) and coherence (upper and lower part of the SFL stimulus walking either in the same or in opposite direction) discrimination, could be performed reliably [2]. In the present work, we developed an ideal-observer-model based on position signals in order to obtain a quantitative grasp on the role of position information in the perception of biological motion. We analyzed model behavior and compared it to experimental data.

2 Methods

2.1 Experiments

For the classical biological motion stimulus, we used an algorithm adapted from Cutting [3]. It computes the joint positions for a point-light display (classical walker) giving the impression of a person walking on a treadmill. For the SFL stimulus, the point-light positions were computed to be somewhere between the

joints, the exact placement changing randomly from one frame to the next. The walker subtended 5 by 11 degrees of visual angle and consisted of white dots. Each animation frame was shown for 52ms. The entire stimulus lasted 2.1s. Ref. [2] provides more detailed information on the stimulus.

In each experiment 2-6 observers participated. They watched the walker stimulus on a dark monitor screen and performed one of several discrimination tasks.

2.2 Simulations

Experimental discrimination tasks were recreated in model simulations. The model used an internal standard of a human walker. We recorded the limb movements of 9 human walkers with a motion tracking system (Ascension MotionStar). A step cycle of the average of these walkers was subdivided into 100 temporally equidistant frames acting as the internal model of the limb configurations of a human walker during a step cycle. For every stimulus frame in the experiment simulation, the model computed the mean distances between the dots in the stimulus frame and the limbs for each frame of the internal standard (figure 2b). The decision for every stimulus frame was then based on the set of standard frames with the minimum distance.

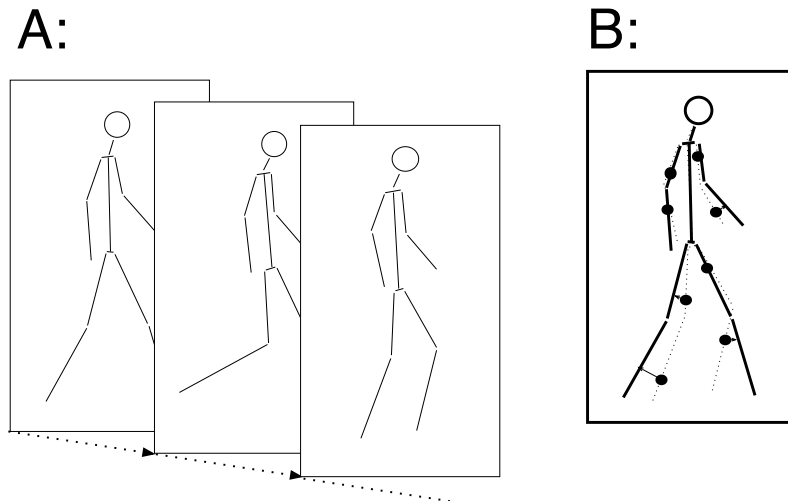


Fig. 2. A: The internal standard consisted of a step cycle of an average human walker subdivided into 100 frames, B: The model's decision is based on linear distance measurements between internal standard and stimulus

In the case of right/left discrimination the model's internal standard consisted of 100 frames of a walker facing and walking to the right and the same number of frames for a walker moving to the left. After the entire stimulus sequence was analyzed, the single answers for each stimulus frame were averaged

to yield an over-all decision. The same approach was taken in the case of coherent/incoherent discrimination, the only difference being that the model's internal walker was subdivided into upper and lower part of the body, a left/right decision was made for each part separately, and then the two decisions were compared for coherence. In both tasks, the model's decisions were therefore based entirely on position information and did not include apparent motion signals between frames.

In the model, we must take into consideration that because of visible persistence [7] for frame durations smaller than 100 ms the number of point-lights perceived at any moment in time is more than the number shown on the display. For instance, for 52 ms frame duration the number of points perceived is about twice the number of dots presented in one frame. To mimic the effect of visible persistence, the model always superimposed any individual frame with the immediately preceding one.

3 Results

3.1 Influence of number of points

As a first quantitative determinant of form information we varied the number of points per frame in several 2AFC tasks. In the direction task, model and human observers had to judge whether the SFL-walker was facing to the right or to the left. In the coherence task they had to discriminate between a coherent and an incoherent walker. A step cycle of the stimulus consisted of 40 frames with a duration of 52 ms (5 monitor refreshes) each. Figure 3a,b shows that the

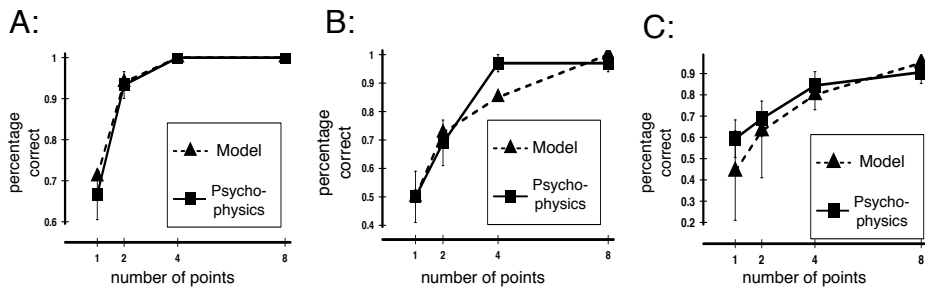


Fig. 3. Comparison of correct answers between model and psychophysical data for A: right/left - , B: coherent/incoherent - and C: forward/backward discrimination

percentage of correct answers increased with rising number of points, both for the model and for the human observers. The similarity between model and human data is surprising as the model does not use any information about the local motion of the points nor about the sequence of the frames. This suggests that the major information used by human observers in the direction and coherence

tasks is frame-by-frame position information, rather than motion signals derived from an analysis of the frame sequence.

We next wished to study a task which cannot rely on form information alone, but which requires sequence analysis. Therefore, we asked observers in a further experiment to discriminate a forward moving display from a backwards moving display. This required the analysis of temporal order over animation frames. The model computed again the distance measures for individual frames but thereafter took the temporal order of the frames into account. Again, performance strongly depended on the number of points per frame (figure 3c). However, the slope was not as steep as for the two previous tasks and performance did not reach 100 percent. Nevertheless, model and psychophysics were again strikingly similar.

3.2 Influence of point lifetime

Beintema and Lappe [2] investigated the potential contribution of local motion signals by prolonging the time over which each light point stayed at one position before jumping to another position (52, 104, 208, or 416 ms) in the direction discrimination task. They argued that if local motion contributes to the perception of biological motion one would expect the percentage of correct answers to increase with prolonged lifetime. But instead of an increase the performance remained constant or showed even a slight decrease with longer lifetimes. Beintema and Lappe speculated that perhaps the reduction in the number of independent position samples that resulted from the increased lifetime led to the decrease in performance.

Model simulations supported this hypothesis (figure 4) as they revealed the same qualitative and quantitative behavior as psychophysical data. This confirms that human observers do not take advantage of additional motion signals. Instead the reduced position information leads to a decline in correct perception rate.

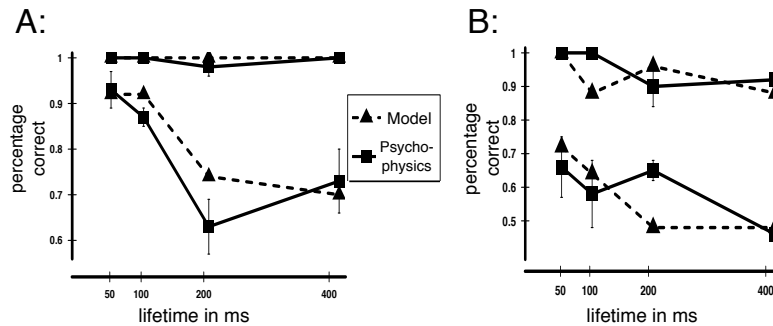


Fig. 4. The influence of lifetime on the percentage of correct answers for A: 8 (two upper curves) and 2 (two lower curves) points and B: for 4 (two upper curves) and 1 (two lower curves) point. Comparison between model and psychophysics

Beintema and Lappe [8] also investigated the potential contribution of local motion signals in the forward/backward discrimination task. In this task, too, prolonged point lifetime did not aid performance. Model simulations showed again similar behavior. No positive influence of prolonged lifetime on the correct answers was observed. This strengthens the conclusion that motion signals do not contribute to performance in this task.

4 Summary and discussion

We investigated the role of position signals in the perception of biological motion using a novel biological motion stimulus that allowed to vary the availability of motion signals. We compared psychophysical studies with an ideal-observer-model that relied only on position information. All experiments revealed striking similarities between model and human data. This suggests that perception is possible from the analysis of form information alone. The model demonstrated that two common psychophysical tasks, direction discrimination and coherence discrimination, could be solved with the same accuracy as human observers without using any motion information. A further task, the discrimination between forward and backward display of a walking person, clearly involved a judgment of motion direction. The model was able to solve this task with the same accuracy as human observers by first analyzing static postures of single frames and then the order of frames in the sequence. Thus, also in this case visual motion signals were not needed.

6 Literatur

1. G. Johansson. Visual perception of biological motion and a model for its analysis. *Percep.Psychophys.*, 14:201–211, 1973.
2. J. A. Beintema and M. Lappe. Perception of biological motion without local image motion. *Proc.Nat.Acad.Sci.USA*, 99:5661–5663, 2002.
3. J. E. Cutting. A program to generate synthetic walkers as dynamic point–light displays. *Behav.Res.Meth.Instrumentation*, 10:91–94, 1978.
4. J. E. Cutting. Coding theory adapted to gait perception. *J.Exp.Psychol.: Hum.Percept.Perform.*, 7:71–87, 1981.
5. L. Vaina, M. Lemay, D. C. Bienfang, A. Y. Choi, and K. Nakayama. Intact biological motion and structure from motion perception in a patient with impaired motion mechanisms: A case study. *Vis.Neurosci.*, 5:353–369, 1990.
6. P. McLeod, W. Dittrich, J. Driver, D. Perrett, and J. Zihl. Preserved and impaired detection of structure from motion by a "motion-blind" patient. *Vis.Cognition*, 3:363–391, 1996.
7. M. Coltheart. Iconic memory and visible persistence. *Percep.Psychophys.*, 27:183–228, 1980.
8. J. A. Beintema and M. Lappe. The role of local position and motion signals in biological motion perception. *Perception*, 30 (suppl.), 2001.