
Independence of contour and biological-motion cues for motion-defined animal shapes

Anne Bellefeuille, Jocelyn Faubert[¶]

Ecole d'Optométrie and Département de Psychologie, Université de Montréal, Montréal, Quebec, H3C 3J7, Canada; e-mail: jocelyn.faubert@umontreal.ca

Received 4 September 1996, in revised form 25 October 1997

Abstract. The effects of different kinds of cues on the perception of second-order motion-defined animal shapes were assessed. In the first experiment discrimination thresholds for motion-defined animals without biological motion (non-BioM) were compared with motion-defined animals with biological motion (BioM). The results show no significant difference between the two conditions, suggesting that BioM does not interact with simple contour motion. In order to isolate the relative strength and interaction between the motion cues a second experiment was conducted where four conditions were used. The first condition consisted of animal contours with non-BioM, the second condition consisted of animal contours with BioM, the third condition was composed of dots present at the joints of the animals with non-BioM, and the fourth condition was composed of dots with BioM. In all cases the animal shapes traveled across the screen for a given number of frames. As in the first experiment, the results of the second study show no interaction between cues. Furthermore, the data show that the thresholds are similar whether BioM or contour cues are presented. The only condition which is significantly different is the condition without either contour or BioM cues. It is concluded that the form representation generated from these cues in motion-defined animal shapes consists of separate mechanisms which appear equally efficient for discrimination and which do not interact with one another.

1 Introduction

The visual system can interpret impoverished images of characteristic locomotion patterns. The cues provided by this characteristic motion pattern have been labeled biological-motion (BioM) cues. In the original study, Johansson (1973) attached small light sources to the points of articulation of a walking person and presented this display in darkness to remove all other visual information. He found that subjects had a vivid impression of a walking human being. Following this original demonstration, many involving using the same kind of display have shown that the perception of BioM patterns could be used to discriminate gender and identity of familiar individuals (Cutting 1978; Cutting and Kozlowski 1977; Cutting et al 1978; Kozlowski and Cutting 1977), walking direction (forward or backward) (Mather et al 1992), ambulating mode (Bertenthal et al 1985; Fox and McDaniel 1982; Jansson and Johansson 1973), and even sign-language interpretation (Poizner et al 1981). Mather and West (1993) explored whether subjects could also recognize animals on the basis of their BioM patterns. They animated stop-action photographs of various animals, recreating the typical motion pattern characterizing each animal. They added to each frame a series of bright dots placed on the joints of the animals and presented this animated sequence of dots on a computer screen. Subjects could identify the animals when presented the animated pattern (dynamic display) but not under a static condition (presentation of a single frame). They concluded that the ability to interpret BioM was not restricted to human movements but generalized to animals.

To date, all studies exploring BioM have involved luminance-defined stimuli, ie two or more areas of the images varied in their mean luminance. These kind of stimuli have been described as first order (Cavanagh and Mather 1989) and are known to

[¶] Author to whom all correspondence should be addressed.

produce a robust impression of motion. However, two areas of an image can have the same mean luminance but differ by second-order properties such as texture, motion, or binocular disparity. In fact, the visual system can analyze and interpret stimuli defined by any of these attributes. Stimuli can be defined by luminance, color, texture, stereopsis, and movement (Cavanagh 1988). The analysis of the different attributes would presumably be achieved by different functional pathways and, possibly, different areas of the brain. Luminance is thought to be one of the most salient and powerful attributes defining contours (Marr 1980). Our first question was whether the visual system could interpret a pattern of BioM when the stimulus is defined solely by motion, with no difference in luminance.

Humans, and probably animals, use different cues provided by the environment for recognition. The shape or contour even when incomplete is a powerful cue to recognize any stimulus. Furthermore, shape can be seen even when contour boundaries are invisible or camouflaged, as evidenced by the attribute-defined stimuli and the BioM studies mentioned earlier. However, the relative strength of BioM for recognition is still unknown. Furthermore, is recognition facilitated by the addition of BioM when contour information is present? Intuitively, we hypothesize that if more cues are available, such as contour and BioM, then recognition should be achieved with less information. If both types of information are processed by a common system, then we would achieve recognition in more impoverished situations on the basis of probability summation. Neurophysiological findings on BioM stimuli and contour stimuli suggest that they may share the same neuronal structure (Bruce et al 1981; Desimone et al 1984; Oram and Perrett 1994; Perrett et al 1985). There is, however, neuropsychological evidence that there may be a specialized system for action recognition which could be independent (Lassonde et al 1993; Vaina et al 1990).

Experiment 1 served to answer two questions: whether the visual system could interpret BioM patterns when they were solely motion defined (i.e. no difference in luminance) and whether recognition from BioM and contour cues were independent. In experiment 1, we assessed whether the addition of BioM to an animal shape defined by moving contours would lower the discrimination threshold. In experiment 2 we isolated the contour and BioM cues to assess their relative strength by testing four conditions where BioM was either present or absent while contour information was either present or absent.

2 Experiment 1: Addition of biological motion to motion-defined animal shapes

The purpose of this experiment was to determine whether the addition of realistic BioM patterns could facilitate the recognition of forms defined by second-order motion.

2.1 Method

2.1.1 Subjects. Eight subjects (23 to 30 years of age) with normal or corrected-to-normal vision participated.

2.1.2 Stimuli and apparatus. Sequences representing the movements of three animals (a horse, a kangaroo, and a greyhound dog) were selected from Macromind Director@. Five to eight images of each animal were chosen to represent characteristic motion. The images were subsequently modified as follows. The size of the animals was normalized so their bodies would cover approximately the same number of degrees of visual angle. Two sizes of each animal were created to control for size-cue discrimination. Large animals covered 3 deg and smaller animals covered 2.25 deg. The complete shape (including extremities) ranged from 4 to 9 cm in height and from 8 to 14.5 cm in length. The areal composition was transformed into a black random-dot pattern and seven levels of coherence for reversal of polarity (50%, 25%, 12%, 9%, 6%, 3%, and 1%) were used (see figure 1).

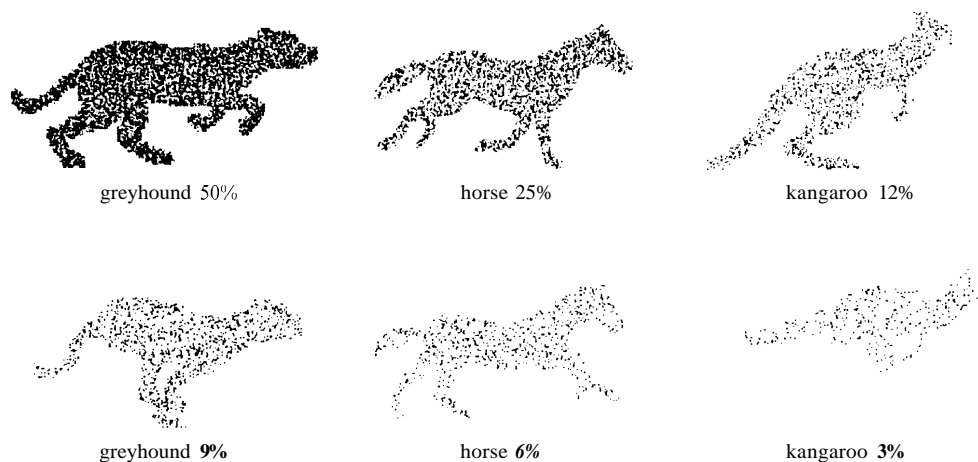


Figure 1. Examples of the three animals used in experiment 1 at different coherence levels. Coherence levels correspond to the percentage of pixels which reverse polarity when the animal travels across the screen. See text for details.

The stimulus background was composed of 50% black and 50% white random dots. Each dot consisted of a single pixel (see figure 2). In order to distinguish the stimuli, sixteen frames were presented in sequence (see figure 3). As the images were presented, the random dots in the area defined by the form reversed polarity. The total dot density of the image thus remained constant, maintaining the mean luminance at 35 cd m^{-2} . For example, in the case of 9% coherence, 9% of the dots reversed polarity, ie black dots became white and white dots became black. Therefore, for any single frame or a static presentation, the foreground (ie the animal) was indistinguishable from the background. Each of the sixteen frames was presented for 0.045 s for a total stimulus duration of 0.72 s , creating a speed of motion of approximately 41 cm s^{-1} (41 deg s^{-1}). There was a partial overlap of the frames.

This type of stimulus manipulation generates a second-order motion stimulus as illustrated in the space–time plot shown in figure 4. Any portion of the animal which covers the background causes reversing polarity changes demonstrated in figure 4. If one were to choose any given column of pixels contained in the animal shape, the space–time plot reveals that there is no consistent and predictable polarity change from one frame to the next. Therefore, information must be integrated over a number of frames to determine the direction and form of the animal. Furthermore, as the coherence level decreases, the average distance between the dots which reverse polarity increases so that the area required for integration of successive frames also increases. The average thresholds for most conditions we tested varied between 10% and 20%

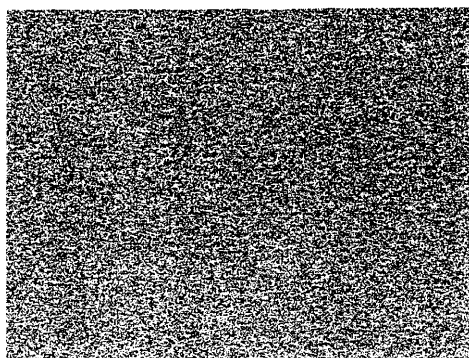


Figure 2. Example of the stimulus background used in the experiments.

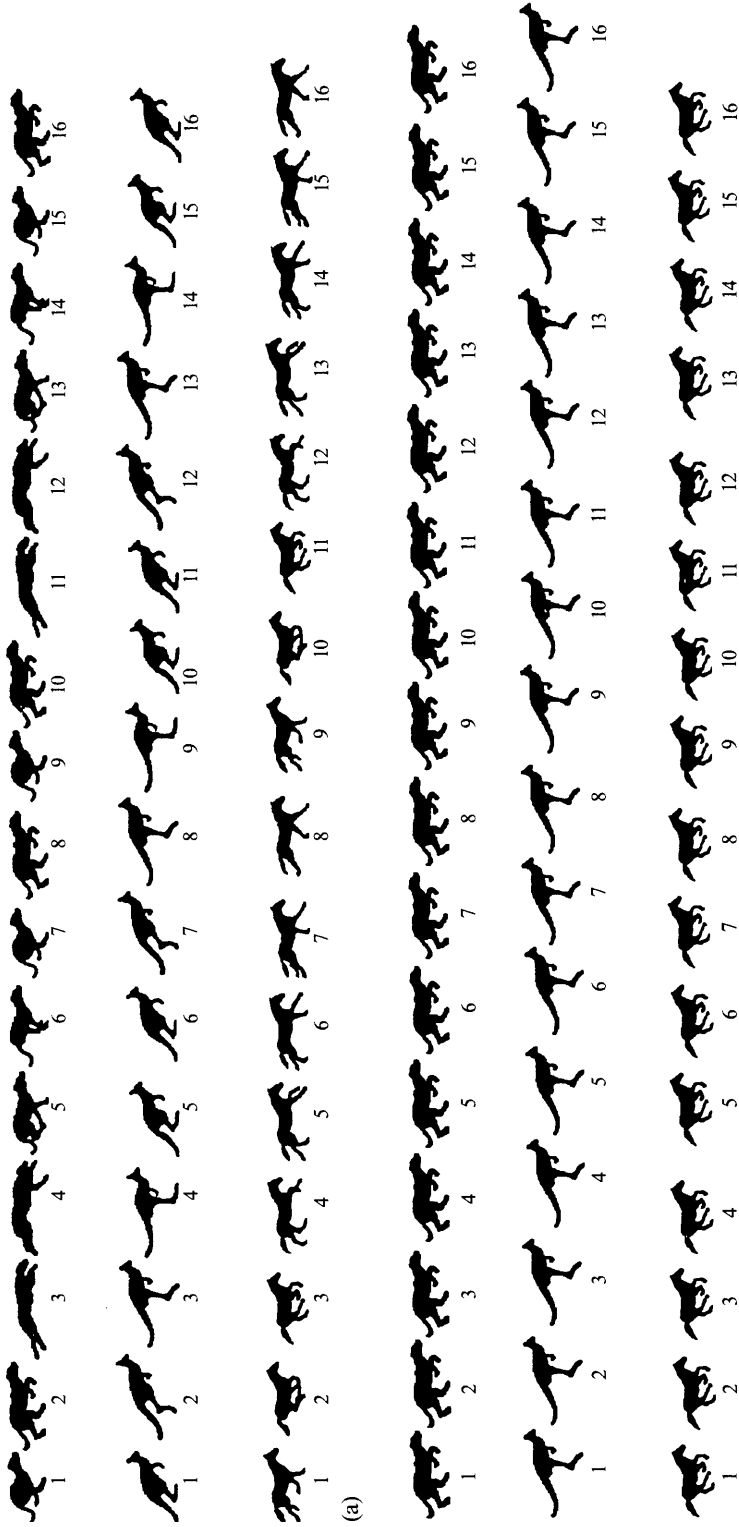


Figure 3. (a) The sequences generated for each animal which created the sensation of BioM (frame 1 – frame 16). One could get a sense of the motion generated by sliding a small window over the patterns in the sequence. (b) The sequences (frames 1 – 16) generated for each animal in the non-BioM condition.

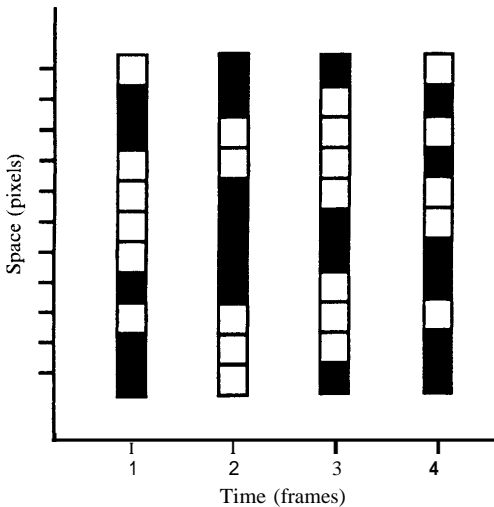


Figure 4. A space–time plot showing the change in polarity for the edge of the animals (or any other part) when it covers the background. The y-axis shows an example of the pixels in a given column of pixels inside the manipulated area of the animal shape and the x-axis represents a sequence of successive frames. One can see that there is no possible correlation between the individual pixel polarity change over time. In experiment 1, it is the entire animal shape which is manipulated this way while in experiment 2 it is only the outline or the joints, according to the condition, while the rest of the animal shape is left unchanged relative to the background.

coherence, meaning subjects were able to discriminate the forms when only 10% to 20% of the dots reversed polarity.

Two motion conditions were created. For the ‘biological-motion’ (BioM) condition, the series of frames representing a motion sequence were cycled, with a total of sixteen frames presented in a single trial sequence (see figure 3a). For the ‘nonbiological-motion’ (non-BioM) condition, one frame was drifted for sixteen consecutive presentations (see figure 3b).

Animations were presented on a standard 13 inch Macintosh RGB monitor interfaced with a Macintosh IICI computer. Stimuli were presented at a distance of 57 cm in a dimly lit room.

2.1.3 Design and procedure. Subjects were first shown a demonstration where they learned to discriminate the animals at 100% coherence. To obtain discrimination thresholds, the method of constant stimuli was used. The subjects were instructed to choose between a horse, a greyhound dog, and a kangaroo and register their response on the keyboard. Each coherence level was presented ten or twenty times, for a total of 70 or 140 trials.

2.2 Results

Discrimination thresholds were calculated with a probit analysis (bootstrap; Foster and Bischof 1991), with a 67%-correct-response criterion. The data for one subject were rejected because this subject obtained very high percentages of correct responses even at our lowest coherence level. Group discrimination thresholds for BioM and non-BioM were compared and no significant difference was found between these conditions on a paired Student’s t-test ($t = 1.57$, $p = 0.17$). The BioM and non-BioM conditions produced mean coherence thresholds of 13.39% and 9.07%, respectively.

2.3 Discussion

Animal-silhouette cues are contained in both the BioM and the non-BioM presentations. Our data suggest that at threshold the addition of BioM cues does not enhance the subject’s performance for object recognition when contour information is provided. As we could not separate the saliency of the BioM cue alone, we designed experiment 2 in order to determine the subject’s ability to use BioM in the absence of contour information.

3 Experiment 2: Biological motion vs outline cues

The purpose of this experiment was to compare the relative strength of form recognition for forms defined by second-order motion when only BioM cues are available, when only contour cues are available, or when both these cues are present or absent.

3.1 Method

3.1.1 *Subjects.* Ten subjects (23 to 30 years of age) with normal or corrected-to-normal vision took part. Four of them also participated in experiment 1.

3.1.2 *Stimuli and apparatus.* The motion sequences of the horse and the greyhound dog were selected and modified as described in experiment 1. They were further modified as follows. In one case, the images were redefined leaving only an outline of the figure. In the other case, dots were placed at the animal's limbs and moving parts (two dots on the tail, two dots on the head, one in the center of the body, and three dots on each leg) and the remainder of the silhouette removed. The new figures (outline or dots) were transformed into black random-dot patterns and seven levels of coherence (100%, 50%, 25%, 12%, 6%, 3%, and 1%) were used (see figure 5). The thickness of the outlines (same area) were generated for both conditions. For example, the small-horse sequence in the outline condition contained approximately 8 190 pixels while in the dot condition the sequence contained 8 220 pixels. The same random-dot background was used and the different motion sequences were designed in a similar fashion to those in experiment 1.

Four motion conditions were created, two involving the outline stimuli and two involving the joints (dot) stimuli. For the outline condition with BioM, the series of frames representing a sequence of motion were cycled, with a total of sixteen frames presented for 1 trial (see figure 6a). For the outline condition with non-BioM, one frame

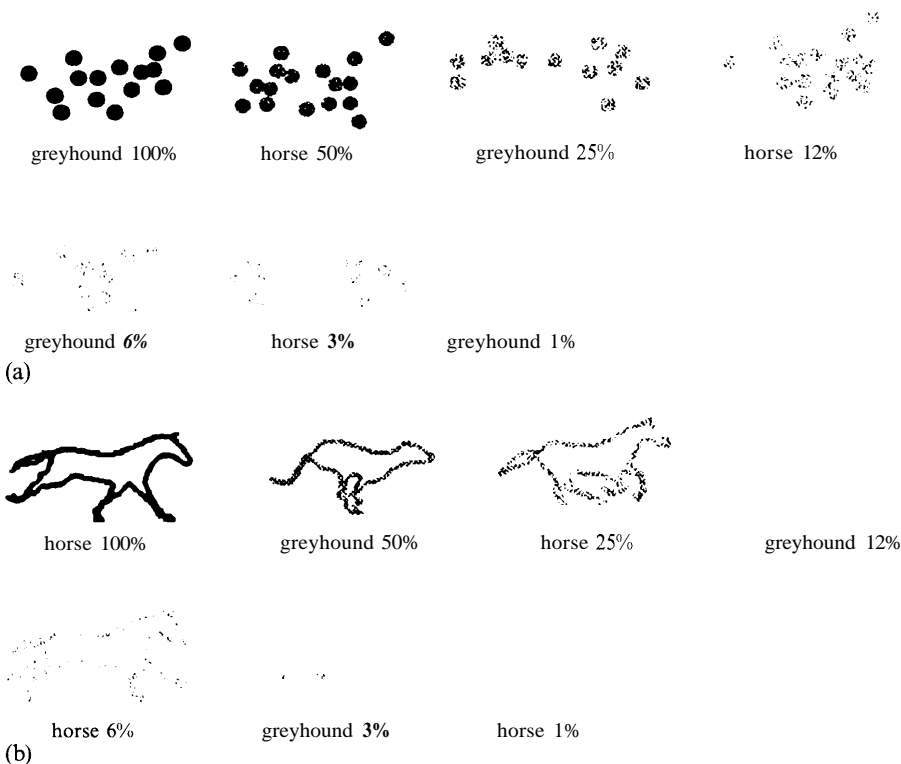


Figure 5. Examples of patterns generated (a) for the dots (joints) and (b) for the contour condition at different coherence levels.

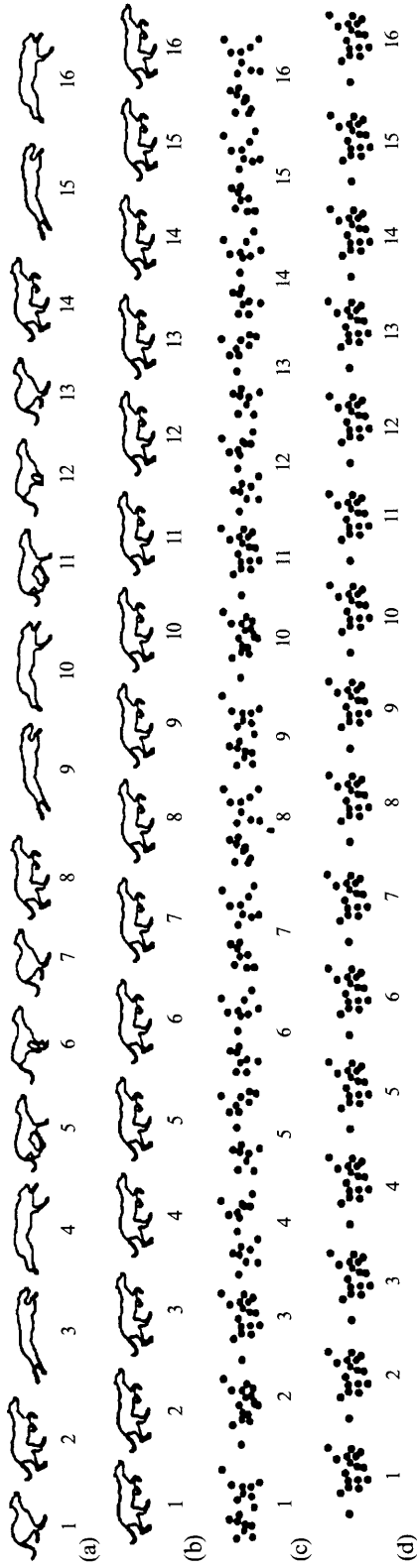


Figure 6. Sequence examples (frames 1 – 16) of the four motion conditions used in experiment 2.

was presented sixteen times (see figure 6b). For the 'joints' condition with BioM, the series of frames representing a sequence of motion were cycled, with a total of sixteen frames presented for 1 trial (figure 6c). For the joints condition with non-BioM, one frame was presented sixteen times (see figure 6d). Animations were presented by using the same specifications as described in experiment 1.

3.1.3 Design and procedure. Subjects were first shown a demonstration where they learned to discriminate the animals at 100% coherence. To obtain discrimination thresholds, the method of constant stimuli was used. Each coherence level was presented twenty times, for a total of 140 trials. The subjects were instructed to identify the presented animal (horse or greyhound).

3.2 Results

Thresholds were determined as in experiment 1, with the exception that threshold was established at 75%. Group discrimination thresholds for the four conditions are presented in figure 7. A 2×2 ANOVA with a posteriori Tukey tests between the treatment conditions was performed on the data. Significant simple main effects were found for contour ($F = 6.72, p = 0.03$) and BioM ($F = 7.78, p = 0.02$). A significant interaction between contour and BioM was also found ($F = 8.01, p = 0.02$). A posteriori Tukey HSD tests revealed that the significance was derived from the difference between the condition where no BioM and no contour information was present (spots drifting) and the remaining three conditions. However, none of these three conditions was significantly different from one another. These results are apparent in figure 7, in which it can be seen that discrimination from the joints-alone condition is much more difficult than the other three conditions. The discrimination threshold was 13.98% for outline with BioM, 13.77% for outline with non-BioM, 17.70% for joints with BioM, and 46.90% coherence for joints with non-BioM.

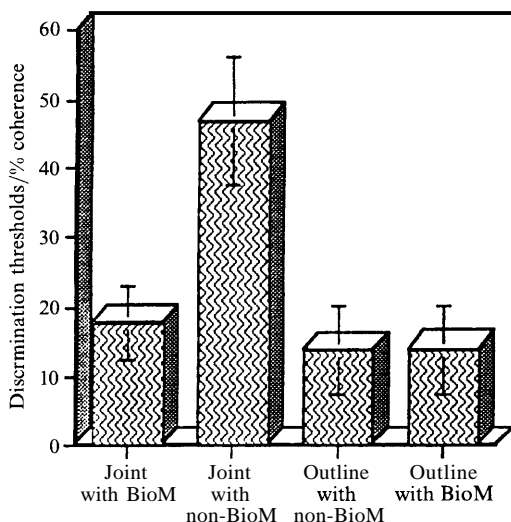


Figure 7. Mean discrimination thresholds for the four conditions tested in experiment 2.

4 General discussion

Previous experiments have shown that subjects are capable of recognizing luminance-defined BioM stimuli. The present study further shows that BioM patterns can also be interpreted when the stimuli are defined by second-order properties.

To determine what aspects of the motion-defined object were important in recognition, we determined discrimination thresholds in the presence or absence of two cues, BioM and continuous contours. Subjects performed poorly when both of these cues

were removed even if motion information remained (drifting joints). The fact that elevated thresholds were obtained when the joints were presented alone without the BioM and contour cues supports Johansson's (1973) idea that it is the pendulum motion and the interactions between joints that produce recognition rather than a linkage system, ie analysis of distance between joints.

When the object contour is present, the subjects appear to discriminate equally well in the presence or absence of BioM. Similarly, when discrimination thresholds for objects with BioM either with (outline) or without (joints) contour were compared, there was no statistically significant difference. These results suggest that there is no summation of the information for recognition because the subjects do not perform better when both cues are present.

With experiment 2, we failed to show a significant difference between the conditions of BioM alone, contour alone, and contour with BioM. The lack of difference between contour information and contour with BioM information was also suggested by the results obtained in experiment 1. What is clear from the results of experiment 2 is that the condition which contains motion but does not have either contour or BioM cues produces significantly higher thresholds than the other three conditions. The difference of percentage coherence thresholds obtained between the other three conditions is very small; 3.718% between the BioM-with-no-contour condition and BioM with contour, and 3.928% between the BioM-with-no-contour condition and the condition with non-BioM and contour cues. The magnitude of these differences does not appear to be biologically significant. With our data, we have calculated that we could determine (with 80% power) whether there was a statistically significant difference of percentage coherence thresholds for the three conditions mentioned above if these differences were in the order of magnitude of 20% or more. In other words, we have sufficient power to determine that the biological and contour cues do not summate prior to recognition if the difference generated by this summation was at least 20% or more between the conditions. Our experience with coherence motion thresholds is that it is not uncommon to observe differences of this magnitude (20%) between naive subjects tested on such procedures. In the present experiment most of our subjects were naive. Nevertheless, we must be cautious in our conclusions that a summation effect is not possible. This lack of significant difference *suggests* that the contour and BioM cues are independent for animal-form-recognition thresholds but it does not necessarily *determine* that this is the case.

Rivest and Cavanagh (1996) described the localization of contours defined by luminance, color, texture, and motion, and showed interactions between the various pairings. Their results support the notion that the information associated with each kind of contour is integrated at a common site. In contrast, our results show neither summation nor interference of the two types of motion cues and, furthermore, they show equivalence in signal strength of both cues separately. In fact, it was striking that the subjects performed equally well when given BioM alone or contour alone. These findings suggest at least two separate pathways of analysis for form recognition that are independent of one another. One pathway relies on the analysis of the contour and the other relies on BioM characteristics to recognize form. Both lead to form recognition independently. Once the form is identified, a decision can be reached. This idea is schematized in figure 8. If both pathways were not independent, a common site for form recognition would be reached and a summation effect would be found; ie the addition of information from different cues would facilitate recognition. The dotted pattern in figure 8 illustrates the summation model, which does not correspond with our data. This idea of independent systems in recognition was also suggested by Mather and colleagues (Mather et al 1992). They state that "the visual system may rely heavily on detecting such characteristic movement patterns during recognition rather than on constructing a full structured representation of the body" (page 154).

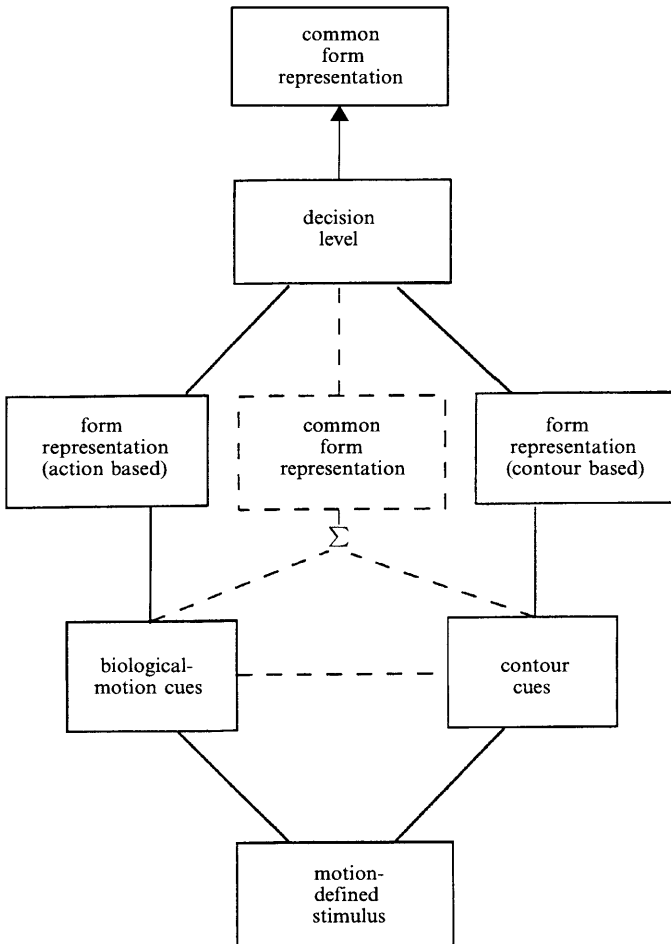


Figure 8. Proposed mechanisms for form recognition of motion-defined animal shapes. The dotted lines show the possible interactions which were hypothesized but not supported by our data.

There is neuropsychological evidence that there may be an independent system for action recognition. Vaina et al (1990) described a patient with impaired motion mechanisms but who could nonetheless perform very well on Johansson's BioM task. Lassonde et al (1993) studied two patients, one with a dorsal-pathway lesion and one with a ventral-pathway lesion, who both had intact global motion sensitivities. The ventral-lesion patient could not identify static images of animals but was extremely good in recognizing animals with a version of the animal-motion task used in the present study. On the same task, the patient with the dorsal lesion could not identify and discriminate any of the animals presented even at 100% coherence but had no difficulty when they were presented in static form. This demonstrates that there are at least two pathways to form representation, one via the dorsal pathway and one via the ventral pathway.

To date, in few studies has the neuronal structure supporting the analysis of form from BioM been examined. Perrett et al (1985) and Oram and Perrett (1994) found cells within the anterior superior temporal polysensory area of the monkey that respond to BioM stimuli. However, cells within this area are also known to respond to static form of the body (Bruce et al 1981; Desimone et al 1984). These physiological findings suggest that BioM stimuli and contour stimuli share the same neuronal structure.

If there are two separate form representations as we suggest (one influenced by contour information and the other by action-based information) one must ask why these two systems do not interact. There is evidence from the study of other types of motion effects that information from very different origins can interact. For instance, Faubert and von Grunau (1995) demonstrated that both high-level and low-level information can interact in the motion-induction illusion. There was indirect evidence of the interaction of contour versus BioM cues in the present study because subjects occasionally reported that contour-defined objects actually appeared to be moving in a biologically correct fashion, or that the joints condition with BioM appeared to have a contour. We are currently investigating these perceptual illusions.

Acknowledgement. This project was partially supported by NSERC grant OGP0121333 to Jocelyn Faubert.

References

- Bertenthal B I, Proffitt D R, Spenter N B, Thomas M A, 1985 "The development of infant sensitivity to biomechanical motions" *Child Development* 56 531–543
- Bruce C J, Desimone R, Gross C G, 1981 "Visual properties of neurons in the polysensory area in superior temporal sulcus of the macaque" *Journal of Neurophysiology* 46 369–384
- Cavanagh P, 1988 "Pathways in early vision", in *Computational Processes in Human Vision: An Interdisciplinary Perspective* Ed. Z Pylyshyn (Norwood, NJ: Ablex) pp 245–289
- Cavanagh P, Mather G, 1989 "Motion: The long and short of it" *Spatial Vision* 4 103–109
- Cutting J, 1978 "Generation of synthetic male and female walkers through manipulation of a biochemical invariant" *Perception* 7 393–405
- Cutting J, Kozlowski L T, 1977 "Recognizing friends from their walk: Gait perception without familiarity cues" *Bulletin of the Psychonomics Society* 9 953–956
- Cutting J E, Proffitt D R, Kozlowski L T, 1978 "A biochemical invariant for gait perception" *Journal of Experimental Psychology* 4 357–372
- Desimone R, Albright T D, Gross C G, Bruce C, 1984 "Stimulus-selective properties of inferior temporal neurons in the macaque" *Journal of Neuroscience* 8 2051–2062
- Faubert J, Grünau M von, 1995 "The effect of two spatially distinct primers and attribute priming on motion induction" *Vision Research* 35 3119–3130
- Foster D, Bischof W, 1991 "Thresholds from psychometric functions: superiority of bootstrap to incremental and probit variance estimators" *Psychological Bulletin* 109 152–159
- Fox R, McDaniel C, 1982 "The perception of biological motion by human infants" *Science* 218 486–487
- Jansson G, Johansson G, 1973 "Visual perception of bending motion" *Perception* 2 321–326
- Johansson G, 1973 "Visual perception of biological motion and model for its analysis" *Perception & Psychophysics* 14 202–211
- Kozlowski L, Cutting J, 1977 "Recognizing the sex of a walker from a dynamic point-light display" *Perception & Psychophysics* 21 575–580
- Lassonde M, Faubert J, Schiavetto A, Bilodeau L, 1993 "Relative contribution of the dorsal and ventral visual systems to motion-defined form" *Investigative Ophthalmology & Visual Science, Supplement* 34 795
- Marr D, 1980 "Visual information processing: the structure and creation of visual representations" *Philosophical Transactions of the Royal Society of London, Series B* 290 199–218
- Mather G, Radford K, West S, 1992 "Low-level visual processing of biological motion" *Proceedings of the Royal Society of London, Series B* 249 149–155
- Mather G, West S, 1993 "Recognition of animal locomotion from dynamic point-light displays" *Perception* 22 759–766
- Oram M W, Perrett D I, 1994 "Responses of anterior superior temporal polysensory (STPa) neurons to 'biological motion' stimuli" *Journal of Cognitive Neuroscience* 6 99–116
- Perrett D I, Chitty A, Mistlin A, Harries M, 1985 "Visual cells sensitive to biological motion" *Behavioural Brain Research* 16 153–170
- Poizner H, Bellugi U, Lutes-Driscoll V, 1981 "Perception of American sign language in dynamic point-light displays" *Journal of Experimental Psychology: Human Perception and Performance* 7 430–440
- Rivest J, Cavanagh P, 1996 "Localizing contour defined by more than one attribute" *Vision Research* 36 53–66
- Vaina M L, Lemay M, Bienfang C D, Choi Y A, Nakayama K, 1990 "Intact 'biological motion' and 'structure from motion' perception in a patient with impaired motion mechanisms: A case study" *Visual Neuroscience* 5 353–369