



## Attention-based visual routines: sprites

Patrick Cavanagh<sup>a,\*</sup>, Angela T. Labianca<sup>a</sup>, Ian M. Thornton<sup>b,1</sup>

<sup>a</sup>*Vision Sciences Laboratory, Department of Psychology, Harvard University, 33 Kirkland Street, Cambridge, MA 02138, USA*

<sup>b</sup>*Cambridge Basic Research, Nissan Research and Development Inc., Cambridge, MA 02142, USA*

Received 12 October 1999; accepted 17 November 2000

---

### Abstract

A central role of visual attention is to generate object descriptions that are not available from early vision. Simple examples are counting elements in a display or deciding whether a dot is inside or outside a closed contour (Ullman, *Cognition* 18 (1984) 97). We are interested in the high-level descriptions of dynamic patterns – the motions that characterize familiar objects undergoing stereotypical action – such as a pencil bouncing on a table top, a butterfly in flight, or a closing door. We examine whether the perception of these action patterns is mediated by attention as a high-level animation or ‘sprite’. We have studied the discrimination of displays made up of simple, rigidly linked sets of points in motion: either pairs of points in orbiting motion or 11 points in biological motion mimicking human walking. We find that discrimination of even the simplest dynamic patterns demands attention. © 2001 Elsevier Science B.V. All rights reserved.

*Keywords:* Attention; Tracking; Motion; Visual search

---

### 1. Introduction

Recognizing an item can call on much more than just an analysis of its static form. Something that moves on a street and makes motor sounds is probably a car or a truck (or maybe a 3-year-old boy). But in addition to characteristic sounds or properties, many objects have characteristic patterns of movement, revealed only over

---

\* Corresponding author. Fax: +1-617-495-3764.

*E-mail address:* patrick@wjh.harvard.edu (P. Cavanagh).

<sup>1</sup> Present address: Max-Planck-Institut fuer Biologische Kybernetik, Spemannstrasse 38, D-72076 Tuebingen, Germany.

some duration of time. Rubber balls bounce on floors, billiard balls bounce less, butterflies dance up and down, Frisbees fly straight, a pencil bouncing off the floor takes an end over end tumble, and doors swing slowly through an arc when opening and bang up against their endpoints. Many of these characteristic patterns of motion are so familiar as to be sufficient for recognition of the object. The case of biological motion is perhaps the strongest evidence for this. A human form is easily recognized from the motions of a set of lights attached to a person filmed while walking in the dark (Johansson, 1973; Neri, Morrone, & Burr, 1998).

How do we accomplish this seemingly effortless recognition of motion patterns? We are not aware of analyzing components of the motion and coming to intermediate decisions. The human walking just seems to pop out of the display. Johansson (1973) proposed that the analysis relied on an automatic and spontaneous extraction of mathematically lawful spatiotemporal relations. But is this act of recognition really effortless? And what about the continued perception of the motion, an analysis which continually adapts our impression of the walker's posture and progress to the moving points in the display. Can that also be as effortless as it seems?

We propose that these characteristic motions are analyzed and interpreted by a special set of operators that we will call 'sprites'. In this paper, we will only address the attentional demands of these operators but we will nevertheless sketch our view of their properties and the role they play. We consider a sprite to be the set of routines that is responsible for detecting the presence of a specific characteristic motion in the input array, for modeling or animating the object's changing configuration as it makes this stereotypical motion, and for filling in the predictable details of the motion over time and in the face of noisy or absent image details. Each different characteristic motion pattern would have its own 'sprite' that would be built up over many exposures to the pattern. These stored recognition and animation routines then allow sparse inputs to support rich dynamic percepts. Many others have stressed that regularities in the world can be captured by efficient, higher-order data structures such as chunks (Miller, 1956), schemata (Bartlett, 1932; Neisser, 1967), frames or scripts (Minsky, 1975; Schank & Abelson, 1977). Characteristic patterns of motion ought to lend themselves quite well to similarly efficient representations.

Separate instances of a characteristic motion are seldom exact repetitions, however. The path of a bouncing pencil can be quite chaotic, depending strongly on many factors (starting position, rotation, surface properties, etc.). The regularities of a bouncing pencil, or a butterfly's flight, or a walking human, lie at a higher level of description of the motion. Significant analysis of the motion pattern must precede any recognition of the regularity and significant computation is then required to use the knowledge of the regularity to predict or animate subsequent motions.

This procedural aspect of a sprite is closely related to the concept of 'visual routines' addressed by Ullman (1984). These routines act on the representations emerging from the initial stage of visual analysis to establish properties and relations that are not explicitly represented in the first stage. Ullman identified elemental processes such as counting, indexing, tracking, and region-filling which could be

organized (compiled) into visual routines to perform a high-level task such as, for example, judging whether a point is inside or outside a complex closed curve.

Importantly, in the original work by Ullman (1984), he assigned the critical operation of shifting the focus of analysis to attentional processes so that his visual routines were exclusively attention-based. In our case, abstracting the high-level description of, say, a bouncing pencil or a point-light walker certainly calls on an analysis of similar or greater complexity than the spatial tasks that Ullman (1984) described. Nevertheless, whether or not these high-level motion descriptors – sprites – require attention remains an open question. It is the central question that we address in this paper.

In the two experiments reported below, we use visual search tasks to examine the attentional load required to perceive a dynamic motion pattern. The displays present one to four motion patterns and the subjects report the presence or absence of a target motion. The relation between reaction time and number of items in the display allows us to evaluate any increase in attentional load with each additional item. If, for example, point-light walkers are recognized effortlessly, then there should be no increase in reaction time as the number of walkers in the display increases from one to four.

Our first experiment examines simple configurations of two moving dots. Although these patterns of motion are relatively simple they are not highly familiar. If attention is required to discriminate between configurations, it may be because we do not have highly efficient routines, sprites, to handle them. Our second experiment examines highly familiar configurations involving human motion. The movements in these stimuli are more complex but extremely familiar. If any dynamic patterns can be discriminated without attentional load, we believe it should be these patterns.

## **2. Experiment 1**

In the first experiment, observers had to discriminate between two different orbital motions. In each stimulus, two lights rotate around each other while moving around a central fixation point. We will describe the two stimuli briefly before examining the similarities between our displays and the classic wheel-generated motions studied in many previous articles (Duncker, 1937; Johansson, 1973; Proffitt, Cutting, & Stier, 1979; Wallach, 1965).

In our first stimulus (Fig. 1a), the motion is like that of a moon orbiting a planet where the planet itself orbits a central ‘star’ (the fixation point). The moon traces out a complex curve (of the cycloid family) around the central point whereas the planet traces a circle. If the smoothly moving light (the ‘planet’) is turned off, the complex nature of the moon’s motion is immediately evident as a looping path of wildly varying velocity. With the smoothly moving ‘planet’ turned on, however, the erratic motion is no longer apparent as the first light is now seen to rotate smoothly at constant velocity around the ‘planet’.

In the second stimulus (Fig. 1b), the two lights rotate around a common center and this central point rotates around the fixation point. In this case, both lights trace

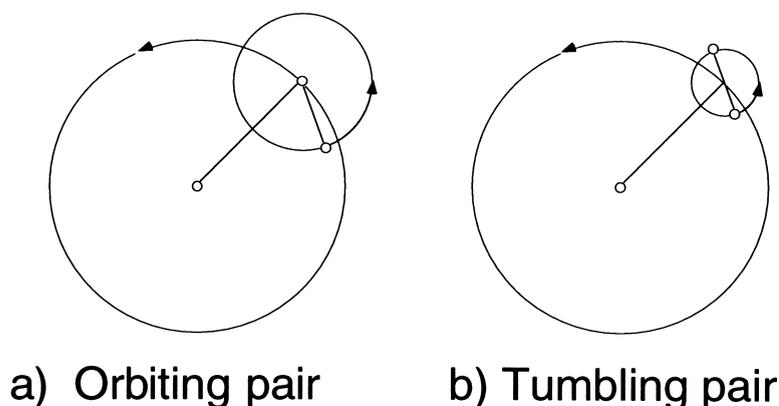


Fig. 1. Trajectories of the orbiting and tumbling motions for dot pairs. Both dot pairs are separated by the same distance and rotate around each other at the same rate. (a) The center of rotation for the orbiting pair is one of the dots so the motion is like that of a moon orbiting a planet which itself orbits the central point. (b) The center of rotation for the tumbling pair is the midpoint between the two dots.

similar intertwined curves (again both are of the cycloid family). Either light alone is seen as moving along a complex path of changing velocity. When the two are present together, however, they create a tumbling motion like that of a baton turning end over end as it moves around the circular path at a constant rate.

The two stimuli have very simple and distinctive motion patterns that might suggest automatic processing of their organization. To examine this question, we asked observers to distinguish between the orbiting and tumbling motions of the two stimuli when there were one or more pairs at different points around the circular path. We consider the tumbling and orbiting motions to be examples of quite familiar motion patterns. Rolling wheels, twirling batons, sticks thrown tumbling through the air, and objects spun around your head on a string fall in one or the other or both of these characteristic motion types. The motions may not have the compelling familiarity of a walking human but they are so simple that we assume that they are good candidates for encoding as characteristic motions or sprites.

Earlier research has looked in depth at a more basic version of these motions: the motions traced out by lights on a wheel rolling along a flat surface. Duncker (1937), for example, noticed that a single light on the rim of a rolling wheel traced out a curve similar to a semicircle (a cycloid). However, when he added a second light to the center of the wheel, the cycloid was no longer seen. Rather, the light on the rim was seen by many observers to rotate around the central light which itself moved in a straight line. Based on the paths of only these two lights (Fig. 2b), the observers 'saw' a wheel in motion even though neither of the motions alone seemed at all wheel-like. Further studies by Johansson (1973), Proffitt et al. (1979), and Wallach (1965) have examined, among other things, the importance of the number and placement of the lights on the wheel in supporting the recovery of the wheel's motion.

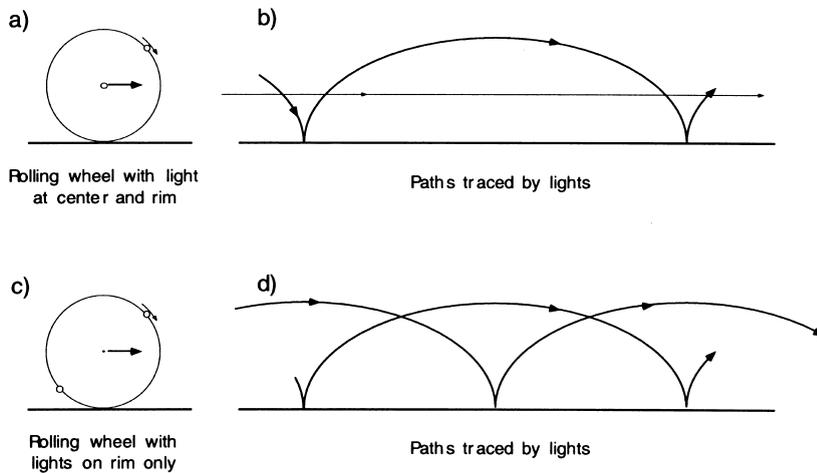


Fig. 2. Trajectories of lights on rolling wheels. (a) One light on the rim and one at the center. (b) Paths followed by the two lights. (c) Two lights on opposite sides of the rim. (d) Paths followed by the lights.

Our first stimulus is equivalent to a translating wheel display with one light on the rim and one in the center (Fig. 2a,b). Our second configuration is equivalent to two lights on the rim at opposite ends of a diameter of the wheel (Fig. 2c,d). According to Proffitt et al. (1979), both of these configurations give rise to equally strong impressions of wheel motion. With displays simulating lights on linearly translating wheels, observers are often drawn to follow the wheel with eye movements. The role of eye movements in recovering the wheel motion has been examined in the traditional wheel displays by Proffitt and Cutting (1979). They found no difference in judgments of wheel motion with or without eye movements. Nevertheless, in our task, we use a central fixation with light pairs orbiting around fixation specifically to avoid eye movements and their effects on reaction time in visual search.

Our two stimuli are constructed to be as similar as possible in terms of basic features and to differ most notably in their high-level organization: tumbling versus orbiting. For example, our tumbling configuration (Fig. 1b) cannot have the same 'wheel' diameter as the orbiting configuration (Fig. 1a). The two lights are on opposite sides of the wheel in the tumbling case but only separated by the radius of the wheel in the orbiting case. If the two 'wheels' had the same diameter, the two lights would be twice as far apart in the tumbling case as in the orbiting case, a difference that would be immediately obvious. To equate the separation of the two lights in both configurations, the radius of the tumbling motion is therefore half that of the orbiting motion. Despite the difference in diameters of the equivalent 'wheels', the rotation of the two lights about their respective midpoints is identical in both stimuli (the midpoints themselves follow different paths).

Our goal in this experiment was to have two easily identified, characteristic motions and determine how quickly each is processed. To do so, we displayed

one, two, three, or four of the dot pairs moving in the same direction at the same rate around the fixation. On a target present trial, one of the pairs is following the target motion, say, tumbling, whereas the others are following the alternative motion, orbiting in this case. We recorded the reaction time to respond that the target was present or absent and analyzed the function relating the reaction time and the number of dot pairs present. If the reaction time did not increase with the number of dot pairs present, we would conclude that the extraction of the characteristic motions calls only minimally on central, attentional resources.

As a control, we also measured the reaction time with only one of the dots present in each pair. The target would then be the smoothly moving ‘planet’ among singleton tumbling dots or one singleton tumbling dot among smoothly moving planets. This gives a measure of the distinctiveness of the single dot motions that make up the characteristic motions of the pairs.

## 2.1. Methods

### 2.1.1. Observers

Twelve paid volunteers, with an average age of about 21 years and with normal or corrected-to-normal vision, participated in this experiment. All participants were naive to the purpose of the experiment and gave informed written consent before the experiments, which were approved by the F.A.S. Human Subjects Committee, Harvard University.

### 2.1.2. Stimuli

The display was presented on a 14 inch 67 Hz Macintosh display driven by a Macintosh 7500/100 programmed in Vision Shell. The two motion patterns we used are shown in Fig. 1. Between one and four of these orbiting or tumbling pairs were presented rotating around the central fixation. The distance from the fixation to the center of rotation of the pair was  $4.3^\circ$  of visual angle at a fixed viewing distance of 57 cm. The center–center separation of the two dots was  $1^\circ$  of visual angle. The dots themselves had a diameter of  $0.4^\circ$  of visual angle. The two dots made one full cycle around each other every 1.5 s while the pair made a full circuit around fixation every 7.5 s. The local and global rotations were always in opposite directions and the direction was set randomly on each trial. The stimuli, when there was more than one, were spaced evenly around the circular path. The initial position of each local rotation was set randomly at the beginning of each trial as was the starting location of the rotation around the circular path. The dots had a luminance of  $70 \text{ cd/m}^2$ , and were presented on a  $20 \text{ cd/m}^2$  background. The temporal onset and offset of the motion pairs on each trial was a step function. The fixation mark was a single static dot at the center of the display. It was identical in size to the moving dots ( $0.4^\circ$  of visual angle in diameter).

### 2.1.3. Procedure

In the control conditions, only one dot of each pair was shown. It was always the smoothly moving dot (the planet) in the orbiting configuration and either of the dots

in the tumbling configuration. In the experimental conditions, both dots of each pair were shown. In half the sessions, an orbiting configuration was the target (with tumblers as distractors) and in the other half, a tumbling configuration was the target (with orbiters as distractors). Across observers, the order of targets and the order of control versus experimental conditions were counterbalanced.

Each set size of one to four motion pairs was presented equally often and half the trials had a target present; half had no target present. The order of presenting set sizes and target present or absent was random. Testing began with four short practice sessions of 16 trials each, followed by the four sessions for the two target types and control or experimental conditions. The order of control and experimental sessions was balanced across observers. Experimental sessions with dot pairs had 64 trials. Control sessions with dot singletons had 32 trials. In all, the testing lasted about 1 h.

Each trial began with a warning tone followed immediately by the presentation of the configurations of dots. Observers had up to 15 s to respond either present or absent by pressing a key. There was no feedback. The reaction time and errors were recorded. The inter-trial interval was 1 s. The observers were told to fixate the central dot at all times. At the start of each session, the target stimulus was identified. Observers were instructed to respond quickly and to avoid making errors.

## *2.2. Results*

The overall error rate was 2.3% in the control conditions and 10.9% in the experimental conditions. The reaction times of correct responses were averaged across observers and are shown in Fig. 3. Linear regression was used to estimate the rate of processing of the moving configurations.

In the control conditions, the single wobbling dot was discovered among smoothly moving dots (Fig. 3a) very rapidly. The search slope for target present trials was about 4 ms per item. Deciding that there was no target present was notably slower with a slope of 173 ms per item. It was harder still to find the one smoothly moving dot among wobbling distractors (slopes of 393 and 658 ms per item for target present and absent, respectively).

In the experimental sessions, finding the tumbling pair among orbiting pairs was extremely slow and finding the orbiting pair among tumbling pairs was the slowest of all. The observers spent almost 1 s or more per item to decide whether it was tumbling or not. It is interesting to consider how long each dot pair was monitored before a decision on the type of motion was reached. For our rough estimate we will assume that search was serial and stopped once the target was found. Since on average only half the pairs need to be checked before finding the target, the slope of 474 ms per pair for the target present trials should be doubled to 948 ms to estimate the average time spent on each pair. The slope for the target absent trials also estimates the time per pair and its value is similar at 1059 ms. The dots complete one rotation around each other every 1.5 s so the approximately 1 s of inspection time is two-thirds of a cycle. Even if we subtract a small portion of that second to allow for attention to shift from pair to pair (Duncan, 1984), it is still the case that the dot pairs require extended monitoring before the representation of their motion is

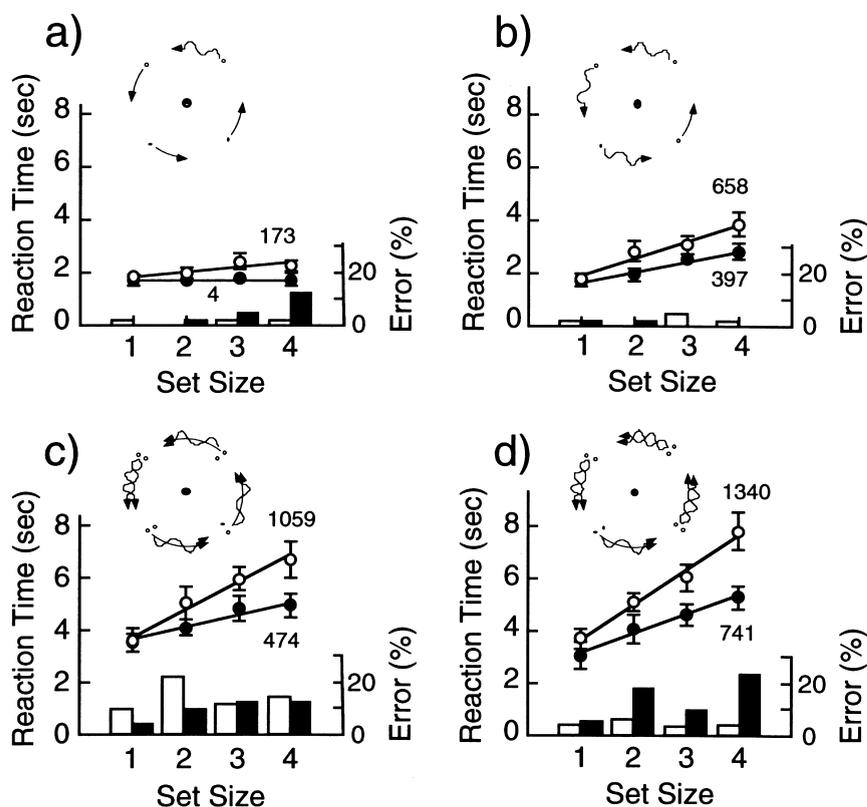


Fig. 3. Reaction time in seconds (left hand vertical axes) as a function of set size for the four search conditions. Target present responses are shown as filled symbols, and target absent responses are shown as outline symbols. Standard errors of the mean are shown as vertical bars when larger than the data symbols ( $\pm 1$  SE). The solid lines show the linear regressions for each data set and the numbers adjacent to the lines are the slopes of the linear regressions in milliseconds per item. Error rates (right hand vertical axes) are shown as histograms at the bottom of each panel with filled bars for target present, and outline bars for target absent trials. (a) Search for a singleton tumbling path among singleton planet paths. (b) Search for a singleton planet path among singleton tumbling paths. (c) Search for a tumbling pair among orbiting pairs. (d) Search for an orbiting pair among tumbling pairs. The data are averages for 12 observers.

settled. Our conclusion is that the perception of these motions is not supported by prepackaged operators specific to their trajectories. The analysis seems to require scrutiny and laborious ‘on-line’ construction of the links and their relative motions. Over many trials of building these constructed motion patterns, performance would undoubtedly improve and a more rapidly engaged operator might emerge that would permit rapid discrimination. No rapid discrimination was evident for the limited exposure offered in our experiments (fewer than 200 trials with dot pairs).

### 3. Experiment 2

In this experiment, we repeated the procedure of the first experiment but now we used point-light walkers, sets of dots in animation which generate a compelling impression of a human walking. Each walker has 11 moving points rather than just the two points of the stimuli in the previous experiment. Although the patterns of motion are also far more complex, they are so familiar that the stimuli are easily and rapidly seen as walking human forms. However, the ease of identifying the motion of a single walker does not tell us much about attentional requirements. The slope of the visual search function could be as steep for these stimuli as the slope for the dot pairs. Conversely, if our intuition about the ease of recognizing these figures is correct, the slopes may be flat – there may be no attentional load.

Thornton, Rensink, and Shiffrar (1999) examined the attention demands of point-light walkers using a dual task method. In the primary task, observers had to report the apparent direction of a point-light walker (leftward or rightward). In the second, observers monitored whether any of a set of rectangles changed orientation. The rectangles appeared at random locations within the same display area as the walker. When the walker was presented in a complex noise field – one designed to disrupt low-level motion integration – the introduction of the secondary task reduced the performance for biological motion to chance levels. They concluded that the perception of biological motion under some display conditions must be an active process, that is, one that is dependent on the availability of attentional resources (Cavanagh, 1992). Interestingly, when a simpler mask was used, the secondary task had little effect, suggesting that a biological motion stimulus might be processed automatically if it were not too degraded.

The attentional load reported by Thornton et al. (1999) may have resulted from the requirements of analyzing biological motion or it may have been due to the requirements of the task itself. The visual search task allows us to separate the overall task demands from the processing demands of each walker. Each additional walker may increase the attentional load, all the while keeping the task demands the same. We displayed one to four walkers simultaneously. On target present trials, one of the walkers was walking to the right and the distractors were walking to the left (the target and non-target directions were reversed for half of the observers). A control task was also run with walkers as targets among non-walking distractors (similar dot motions that did not appear to be a human walking) (Fig. 4). In all cases, observers maintained fixation at the center of the display and responded as quickly as possible whether or not a target walker was present.

#### 3.1. Methods

##### 3.1.1. Observers

Ten experienced observers, members of the Vision Sciences Laboratory, with normal or corrected-to-normal vision, participated in this experiment without reimbursement. All participants were naive to the purpose of the experiment and gave

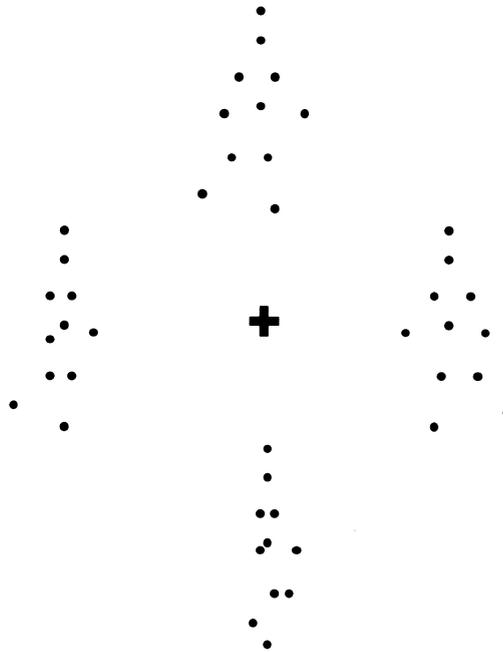


Fig. 4. A static frame from a display with four walkers. This depicts a trial where the target is a figure walking to the left, in this case, the figure at 3 o'clock.

informed written consent before the experiments, which were approved by the F.A.S. Human Subjects Committee, Harvard University.

### 3.1.2. Stimuli

The details of the displays were the same as in Experiment 1 with the following exceptions. The biological motion configuration was generated by modifying Cutting's classic point-light walker algorithm (Cutting, 1978). The set of 11 dots simulated a walker seen in profile with lights on the head, near shoulder, both elbows, both wrists, near hip, both knees and both ankles. The dots were always visible, and they did not disappear when they would be occluded by the walker's body. The walker did not move across the screen but walked in place with either left- or rightward gait. The distance from the fixation to the center dot of the walker subtended about  $4^\circ$  of visual angle, as did the height of the walker. The maximum stride width of a walker was about  $2^\circ$  of visual angle. The dots themselves had a diameter of  $0.2^\circ$  of visual angle. The walker's stride cycle took about 1.3 s, falling within the range of 0.8–2 s per stride reported for normal human walking (Inman, Ralston, & Todd, 1981). The walker's starting phase in its stride and position around the fixation point was selected randomly on each trial. When more than one walker was displayed, the starting phase of the stride for each was assigned randomly and spaced equally around fixation. The dots had a luminance of  $0.1 \text{ cd/m}^2$ , and were

presented on a  $2.12 \text{ cd/m}^2$  background. The fixation mark was a black cross at the center of the display subtending  $0.5^\circ$  of visual angle.

### 3.1.3. Procedure

The details of the procedures were the same as in Experiment 1 with the following exceptions. The target could be either a rightward gait or a leftward gait while the distractors had the opposite gait. The target was fixed for each observer but counter-balanced across observers. Each observer participated in two sessions of 80 trials. They had up to 5 s to respond either present or absent by pressing a key. No warning tone was used. In all, the testing lasted about 15 min. Sessions began with a few practice trials to familiarize the observers with the stimuli and responses.

### 3.2. Results

The overall error rate was 3.5%. The reaction times of correct responses were averaged across observers and are shown in Fig. 5. Linear regression was used to estimate the rate of processing of the walkers.

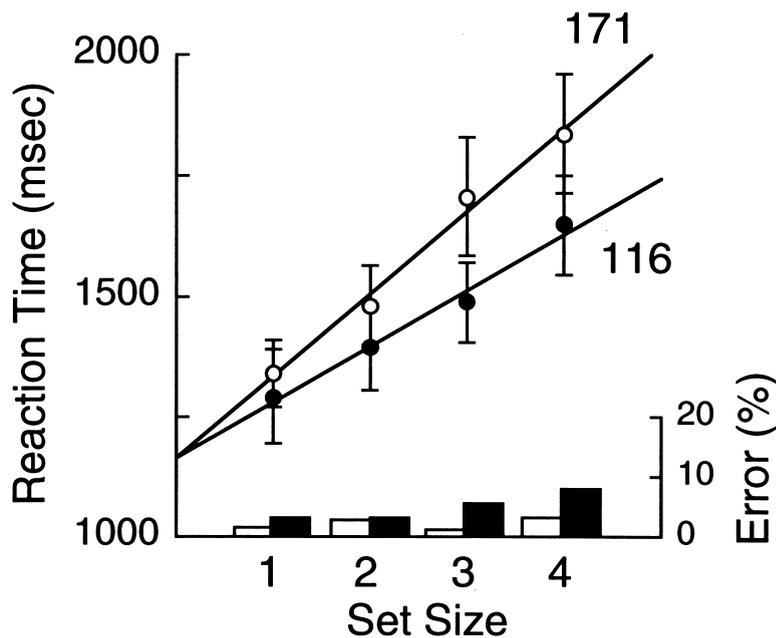


Fig. 5. Reaction time in seconds (left hand vertical axis) as a function of set size for the walker search task. Target present responses are shown as filled symbols, and target absent responses are shown as outline symbols. Standard errors of the mean are shown as vertical bars ( $\pm 1$  SE). The solid lines show the linear regressions for each data set and the numbers adjacent to the lines are the slopes of the linear regressions in milliseconds per item. Error rates (right hand vertical axis) are shown as histograms at the bottom of each panel with filled bars for target present, and outline bars for target absent trials. The data are averages for ten observers.

Finding the leftward gait among rightward distractors (or vice versa) was much faster than finding the tumbling or orbiting target in the previous experiment. Nevertheless, there was a significant slope in excess of 100 ms per walker. Although familiarity helped the search tremendously, search remained serial. There was no pop-out of the odd walker. Again we estimated how long each walker was monitored before a decision on gait direction was reached. The slope of 116 ms per walker is doubled to 232 ms to estimate the average time spent on each walker in the target present trials. The slope for the target absent trials suggests somewhat less at 171 ms per walker. The walkers complete one cycle of their stride every 1.3 s so even 232 ms is only one-fifth of a cycle. Once any time required for shifting attention from walker to walker is subtracted, it appears that even though the walkers must be scrutinized serially, the direction of the gait is determined very rapidly.

We were interested in whether left–right confusions might have contributed to the difficulty of the search. We ran a control condition with a standard walker as the target among impossible walkers (the dots comprising one arm and one leg were phase shifted relative to the rest of body, providing a very distinctive, non-rigid, non-human, skipping or dancing gait). The search rate here, averaged across four subjects, was even slower than for the left versus right task with slopes of about 160 ms per item (and higher still, 230 ms per item, for a non-walker target in walker distractors).

#### **4. Conclusions**

Distinguishing the tumbling from the orbiting motions in Experiment 1 was effortful. When several dot pairs were present, they all appeared to be tumbling. Only by attending to an orbiting pair did it become obvious that one of the dots was moving smoothly around the central dot and that the other was orbiting the first. The tumbling percept seemed to be the default that was seen with minimal attention for both of the motions; extracting the true relative motion required extended scrutiny. The estimated processing time per dot pair was about 1 s or more indicating that substantial processing was required.

Why were the dot pairs so difficult? Recall that in the control conditions a single tumbling target could be easily picked out from individual smoothly moving dots (without their orbiting moons). The highly distinctive motion of the individual dots was totally lost when an additional dot was added. It appeared that only the relative motion of the two dots was available and in terms of relative motion, both the tumbling and the orbiting pairs were constructed to be very similar. To distinguish the two, the observer must attend to the three-dot configuration including the central dot. Only then it is clear that one of the dots is always the same distance from the central point. This very slow process suggests that the subjects have no specialized analyzers that can rapidly identify these motion patterns – no sprites for tumbling or orbiting motions. The trajectories must be individually traced with attention to determine which pattern is governing each dot pair.

Moreover, it appears to be the dynamic aspect of the trajectory that makes it so

difficult. Motion takes place over time and considerable time appears to be required to establish the trajectory of each pair. The static versions of the trajectories of our two motion patterns are much easier to distinguish. The orbiting pattern traces out something like a dollar sign and the tumbling pair something like a figure 8. When we simply presented these short trajectories as spatial patterns all at once, the discrimination was much more rapid (slope of less than 100 ms/item, averaged over four subjects). What makes the real motion case so difficult is that the configuration of dot motions needs to be made explicit, linking each dot to the next and then to the central dot. This is the only way to discover if one of the dots is actually maintaining a constant distance from the central dot.

These articulated links of constant length in tumbling and orbiting motion also form the basis of the structure for the point-light walkers where the light at each joint is separated from the next by a fixed distance. However, for the walkers in Experiment 2, the processing time was much faster at about 200 ms per walker (compared to about 1 s per dot pair). This rate is comparable to some estimates of the dwell time of attention (Duncan, 1984). This suggests that attention was required to notice the gait of each walker in turn but that little processing was required once each walker had been selected. This is evidence that the analysis of a very familiar motion pattern, despite its complexity, can be very rapid. We suggest that this rapid extraction of the motion pattern is the signature of the ‘sprite’ responsible for recognizing and animating the percept of a walking human form. Despite this rapid extraction, our data also show that only one walker at a time can be analyzed. The search rate was still substantial indicating that the operation of at least the ‘walker’ sprite requires attention.

The dual task results of Thornton et al. (1999) suggested that, in some cases, the perception of biological motion can be automatic. However, a search task is perhaps a more sensitive measure of attentional load because dual task interference only reveals an interaction between the two tasks if the combined load exceeds the available capacity. It cannot differentiate between no attentional demands and any combination that is less than the limit available.

If even familiar patterns of motion require attention to be discriminated, what can be the advantage of the routines that support the perception of the pattern? Clearly, it is the same advantage that is offered by any recognition of a familiar pattern. Once enough of the pattern is acquired to recognize it, the rest can be filled in from memory. Sparse inputs can support rich percepts and in the case of a moving object, filling in implies a prediction of likely motions and tracking them with less data than would be otherwise necessary. These advantages have formed the basis of many theories of perception from schemata and schema theory (Bartlett, 1932; Neisser, 1967) to frames and scripts (Minsky, 1975; Schank & Abelson, 1977).

To conclude, we suggest that the visual system acquires and uses stored motion patterns, sprites, which are characteristic of familiar events or objects: the motion of a wheel, the jump of a fish out of water, the way a pencil bounces on the floor when dropped, and the way a fresh egg does not. We use these stored patterns to recognize and then animate our perception of familiar events. Our experience of these animation routines might suggest that they are effortless but our study here shows that they

are not. We claim that the animations are played out by attentive processes in the same way that we can animate a mental image. In the visual case, the input image data act like set points in the progress of the animation but the animation still requires the support of attentive processes.

### Acknowledgements

Thanks to Pawan Sinha for helpful discussions of this work and to Jane Raymond for sharing her observations on Frisbees and butterflies. This research was supported by grant EY09258 to P.C.

### References

- Bartlett, S. F. (1932). *Remembering: a study in experimental and social psychology*. Cambridge: Cambridge University Press.
- Cavanagh, P. (1992). Attention-based motion perception. *Science*, 257, 1563–1565.
- Cutting, J. E. (1978). A program to generate synthetic walkers as dynamic point-light displays. *Behavioral Research Methods and Instrumentation*, 10, 91–94.
- Duncan, J. (1984). Selective attention and the organization of visual information. *Journal of Experimental Psychology: General*, 113, 501–517.
- Duncker, K. (1937). Induced motion. In W. D. Ellis (Ed.), *A sourcebook of Gestalt psychology*. London: Routledge and Kegan Paul.
- Inman, V. T., Ralston, H., & Todd, J. T. (1981). *Human walking*. Baltimore, MD: Williams & Wilkins.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, 14, 201–211.
- Miller, G. A. (1956). The magical number seven plus or minus two: some limits on our capacity for processing information. *Psychological Review*, 63, 81–97.
- Minsky, M. (1975). A framework for representing knowledge. In P. H. Winston (Ed.), *The psychology of computer vision* (pp. 211–280). New York: McGraw-Hill.
- Neisser, U. (1967). *Cognitive psychology*. New York: Appleton.
- Neri, P., Morrone, M. C., & Burr, D. C. (1998). Seeing biological motion. *Nature*, 395, 894–896.
- Proffitt, D. R., & Cutting, J. E. (1979). Perceiving the centroid of configurations on a rolling wheel. *Perception and Psychophysics*, 25, 389–398.
- Proffitt, D. R., Cutting, J. E., & Stier, D. M. (1979). Perception of wheel-generated motions. *Journal of Experimental Psychology: Human Perception and Performance*, 5, 289–302.
- Schank, R., & Abelson, R. (1977). *Scripts, plans, goals and understanding*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- Thornton, I. M., Rensink, R. A., & Shiffrar, M. (1999). Biological motion processing without attention. *Perception*, 28 (Suppl), 51.
- Ullman, S. (1984). Visual routines. *Cognition*, 18, 97–159.
- Wallach, H. (1965). Visual perception of motion. In G. Keyes (Ed.), *The nature and the art of motion*. New York: George Braziller.