

Responses of Anterior Superior Temporal Polysensory (STPa) Neurons to “Biological Motion” Stimuli

M. W. Oram and D. I. Perrett

University of St. Andrews, Scotland

Abstract

■ Cells have been found in the superior temporal polysensory area (STPa) of the macaque temporal cortex that are selectively responsive to the sight of particular whole body movements (e.g., walking) under normal lighting. These cells typically discriminate the direction of walking and the view of the body (e.g., left profile walking left). We investigated the extent to which these cells are responsive under “biological motion” conditions where the form of the body is defined only by the movement of light patches attached to the points of limb articulation. One-third of the cells (25/72) selective for the form and motion of walking bodies showed sensitivity to the moving light displays. Seven of these cells showed only partial sensitivity to form from motion, in so far as the cells responded more to moving light displays than to moving controls but failed to discriminate body view. These seven cells exhibited directional selectivity. Eighteen cells showed statistical discrimination for both direction of movement and body view under biological motion conditions. Most of these cells showed reduced re-

sponses to the impoverished moving light stimuli compared to full light conditions. The 18 cells were thus sensitive to detailed form information (body view) from the pattern of articulating motion. Cellular processing of the global pattern of articulation was indicated by the observations that none of these cells were found sensitive to movement of individual limbs and that jumbling the pattern of moving limbs reduced response magnitude. A further 10 cells were tested for sensitivity to moving light displays of whole body actions other than walking. Of these cells 5/10 showed selectivity for form displayed by biological motion stimuli that paralleled the selectivity under normal lighting conditions. The cell responses thus provide direct evidence for neural mechanisms computing form from nonrigid motion. The selectivity of the cells was for body view, specific direction, and specific type of body motion presented by moving light displays and is not predicted by many current computational approaches to the extraction of form from motion. ■

INTRODUCTION

In the early 1970s it was found that human subjects could interpret extremely impoverished images of human walking. Small light sources were attached to the points of articulation of a walking person (the shoulders, elbows, wrists, hips, knees, and ankles), then all other visual information removed by presenting the stimulus in darkness. Johansson (1973) found that subjects had no difficulty in identifying the stimulus as representing a person walking. Indeed Johansson reported the effect as being “immediate and compelling.” He referred to this type of stimuli as biological motion stimuli. They have also been referred to as moving light displays.

Subjects can perceive a variety of information from such biological motion stimuli, including the gender and identity of familiar individuals (Cutting, 1978; Cutting & Kozlowski, 1977; Cutting, Proffitt, & Kozlowski, 1978; Cutting, Moore, & Morrison, 1988; Kozlowski & Cutting, 1977), whether the individual walks forward or backward

(Perrett, Harries, Benson, Chitty, & Mistlin, 1990a; Perrett, Harries, Chitty, & Mistlin, 1990b; Mather, Radford, & West, 1992), as well as the mode of ambulation (Jansson & Johansson, 1973; Fox & McDaniel, 1982; Bertenthal, Proffitt, Spenter, & Thomas, 1985) and other actions (e.g., sign language, Poizner, Bellugi, & Lutes-Driscoll, 1981). Thus at the human perceptual level biological motion stimuli can give a great deal of information, not only about the nature of the movements but also the form of the individual that is moving. Despite the rich source of information from such nonrigid motion stimuli little is known about the underlying neuronal mechanisms. The similarity of behavioral performance between human and macaque subjects in processing form from motion (Siegel & Andersen, 1990) suggests that the macaque is a suitable model for investigating the underlying neural mechanisms of form from motion processing. In this paper we present the first quantitative analysis of neuronal populations in the macaque monkey that might support the analysis of form from biological motion.

Form and Motion Pathways

It has been suggested that processing of visual information in primates follows two pathways: the ventral "form" or "what" pathway and the dorsal "motion" or "where" pathway (Ungerleider & Mishkin, 1982; Mishkin, Ungerleider, & Macko, 1983; De Yoe & Van Essen, 1988). These two pathways involve several brain areas (Felleman & Van Essen, 1991; Young, 1992). The ventral pathway passes through the areas V1, V2, V4, into posterior, central, and anterior inferotemporal cortex (PIT, CIT, AIT) and the anterior sections of superior temporal sulcus (including area STPa). The dorsal or "motion" pathway flows from V1 through V2, the middle temporal area (MT), also known as V5, and the lateral and dorsal medial superior temporal areas (MSTl and MSTd) and then passes to the frontal eye fields and parietal cortex. The termination areas of this pathway have led to the suggestion that it is involved in the control of eye movements and visuomotor interactions with objects (Goodale & Milner, 1992). The two pathways are not completely separate: outputs from areas MSTl and MSTd also pass through the fundus of the superior temporal sulcus (FST) to the posterior and anterior sections of the superior temporal polysensory area (STPp and STPa, Boussaoud, Ungerleider, & Desimone, 1990). Area STPa therefore receives inputs from both the ventral (form) and dorsal (motion) pathways (Felleman & Van Essen, 1991; Young, 1992). In view of this anatomical convergence, it may not be surprising that some neurons in area STPa show selectivity both for the form and the direction of motion of objects. Single cells in macaque STPa (and more generally throughout the anterior sections of the superior temporal sulcus, STS) have been found to be selectively responsive to the sight of various body movements including walking and articulation of individual limbs (Brothers & King, 1992; Desimone, Albright, Gross, & Bruce, 1984; Hasselmo, Rolls, Baylis, & Nalwa, 1989; Perrett et al., 1990b) and hand actions (e.g., tearing, object manipulation, Perrett, Mistlin, Harries, & Chitty, 1989a; Perrett, Harries, Bevan, Thomas, Benson, Mistlin, Chitty, Hietanen, & Ortega, 1989b). We report here a study of the responses of cells to whole body motion defined under biological motion.

Mechanisms of Sensitivity to Form and Motion

There are three broad possible categories of mechanism by which conjoint selectivity to form and motion could be achieved in the STPa. (1) STPa cells could integrate information about the direction of overall displacement during movement (from the dorsal inputs) and information about the form of the stimulus (from the ventral inputs). (2) Sufficient motion information might be available to STPa cells (from the dorsal route alone) to estab-

lish sensitivity to the patterns of articulation. (3) Selectivity for body movement could be established by combining inputs from multiple cells (in inferotemporal cortex, IT, or STPa) each selective for the same body form but at slightly different spatial positions. Motion sensitivity could in this case derive from inputs from the ventral route alone using circuitry analogous to that proposed for other systems (Barlow & Levick, 1965; Torre & Poggio, 1978). This latter mechanism is the least likely, since cells in the STPa and IT have very large receptive fields (Bruce, Desimone, & Gross, 1981) though changes in sensitivity to stimuli at different positions within the large receptive fields (Gross, 1992) could perhaps be used.

Selectivity that could be used in all three processing schemes has already been documented. Cells selective for the static form of the head and body are found within the STPa and IT (Bruce et al., 1981; Desimone et al., 1984; Hasselmo et al., 1989; Perrett, Rolls, & Caan, 1982; Perrett, Smith, Potter, Mistlin, Head, Milner, & Jeeves, 1984, 1985a; Perrett, Oram, Harries, Bevan, Hietanen, Benson, & Thomas, 1991; Perrett, Hietanen, Oram, & Benson, 1992) as are cells selective for direction of motion but lacking form sensitivity (Gross, Rocha-Miranda, & Bender, 1972; Bruce et al., 1981; Perrett et al., 1985b; Hikosaka, Iwai, Saito, & Tanaka, 1988; Hietanen & Perrett, 1993; Oram, Perrett, & Hietanen, 1993). Utilization of these cell types could support scheme (1). Inputs to the STP from MT and MST are likely to convey motion information but relatively little form information. Area MT and MST contain increasing numbers of cells selective for the direction of motion independent of local contour motion (Albright, 1984; Albright, Desimone, & Gross, 1984; Duffy & Wurtz, 1991a,b; Komatsu & Wurtz, 1988a,b; Mikami, Newsome, & Wurtz, 1986a,b; Rodman & Albright, 1989; Saito et al., 1986, 1989; Snowden et al., 1991, 1992; Tanaka et al., 1986, 1989; Tanaka & Saito, 1989; Zeki, 1974). These inputs could support scheme (2) or be used in conjunction with inputs from cells selective for static form as in scheme (1). Furthermore, IT, which projects to the STPa (Felleman & Van Essen, 1991; Young, 1992), contains cells that are selective for static form and perspective view of the head (Tanaka et al., 1991; Hasselmo et al., 1989; Young & Yamane, 1992). These cells (and also cells within area STPa selective for form) could be used to support scheme (3) or be combined with inputs from cells selective for motion to support scheme (1). Therefore any of the proposed schemes of processing to detect walking bodies under normal lighting could in principle be implemented by cells in the STPa, using either inputs from cells within area STPa or inputs from cells in areas IT and MT/MST. We stress that under scheme (1), the suggested motion input carries only overall translation information. Therefore form information could not be calculated under biological motion conditions since it is only the motion of the light points relative to one another that can be used to extract form-

from-motion (e.g., both left and right profiles moving to the left have the same large field motion signals). Similarly under scheme (3) only overall translation information would be present, so again no sensitivity to biological motion stimuli would be seen. Hence STPa cells would be expected to respond to biological motion stimuli only under the second scheme of processing.

View Specificity

The majority of processing of static form information within the STS and IT cortex appears to be conducted in a view specific manner (Bruce et al., 1981; Desimone et al., 1984; Perrett et al., 1985a, 1991, 1992; Hasselmo et al., 1989). For example, individual cells respond to the left profile view of the head but not the right profile or other views. Such sensitivity to perspective view has also been observed in STPa cells conjointly sensitive to body form and motion; some cells respond selectively to the left profile body view walking to the observer's left (Perrett et al., 1985b, 1990a,b). Neuronal sensitivity to the visual patterns of monkey ambulation in specific directions has been observed in other regions of the macaque temporal lobe (e.g., the amygdala, Brothers & King, 1992). The view sensitivity seen in STPa cells offers an opportunity to quantify sensitivity to form defined by motion, since the response to movement of one body view can be compared to a different view moving in the same direction. Mirror image body views are identical in size, complexity of articulating elements, and angular speed of component movements. Discrimination of responses to different views therefore indicates sophisticated processing of form. Discriminating body view under biological motion conditions has been used in psychophysical tasks to assess quantitatively human perceptual sensitivity to form defined by motion (Cutting et al., 1988; Mather et al., 1992; Perrett et al., 1990a).

The majority of computational models of form from motion in general (Ullman, 1979; Hildreth & Koch, 1987) and biological motion in particular utilize general purpose procedures that are equally applicable to all views of walking bodies and indeed all articulating entities (e.g., Rashid, 1980; Webb & Aggarwal, 1982; Hoffman & Flinchbaugh, 1982; Sugie & Kato, 1987; Sugihara & Sugie, 1984). View sensitivity of cells responsive to body motion is therefore an important attribute to quantify since it is a property that is not predicted on the basis of most current computational approaches to biological motion.

The primary aim of this study was to determine whether cells in the STPa selectively responsive to the sight of walking bodies under normal lighting conditions were sensitive to biological motion versions of the same walking stimuli. Earlier reports suggested that STPa cells might indeed utilize patterns of articulation (Bruce et al., 1981; Perrett et al., 1990a,b) but no systematic study had been made of the extent to which cell sensitivity to body

form and direction of movement was maintained under biological motion conditions.

RESULTS

Cells Selective to Human Walking

From the four subjects, 161 cells were found to be sensitive to walking stimuli out of a total of 6459 cells screened (see Methods). We report here on a subset of these 196 cells (other cells selective for walking stimuli were subjected to studies of tuning for view, direction, and object-part interactions). A total of 72 of the cells found to be selective for the walking stimuli were tested for sensitivity to biological motion (dots and/or the stick figure variation). The selectivity in the responses for human walking could not be attributed to single limb articulation for any of these cells (see Methods). Of the 72 cells selective to walking stimuli, 47 (65%) gave no response above spontaneous activity or control response levels when tested with biological motion stimuli. Thus approximately two-thirds of the cells selective for walking bodies did not show any responsiveness to stimuli where only motion information was available for defining stimulus form. The lack of response indicates the conjoint selectivity shown by these cells, namely that both form and motion information are required to elicit a response (Oram et al., in prep.).

Seven cells (10%) showed a maintained directionality but not view discrimination under biological motion conditions. That is, with moving light displays these cells responded more strongly to both body views moving in the cell's preferred direction than to controls moving in the same direction, spontaneous activity or biological motion in the null (opposite) direction. For instance, in Figure 1 it can be seen that the cell does not maintain the view discrimination seen under normal lighting but responds well to biological motion representations of left and right body views moving to the left. Thus sensitivity to body view was not seen. The cell does, however, maintain direction discrimination. More importantly, responses to biological motion stimuli moving left were greater than responses to controls moving in the same direction, indicating partial sensitivity to body form. The responses to the overall direction of the biological motion stimulus did show clear discrimination between movement to the left and right. No cells were found with the converse selectivity, that is, showing view selectivity but not directional selectivity under moving light displays.

The remaining 18 cells out of the 72 tested (25%) showed full selectivity for biological motion of dot and stick figures. Full selectivity to biological motion is defined here as selectivity for both form (body view) and direction of motion. Ten of these cells were sensitive to moving dot stimuli and eight to stick figure stimuli. Four cells were tested with both moving dot and stick stimuli.

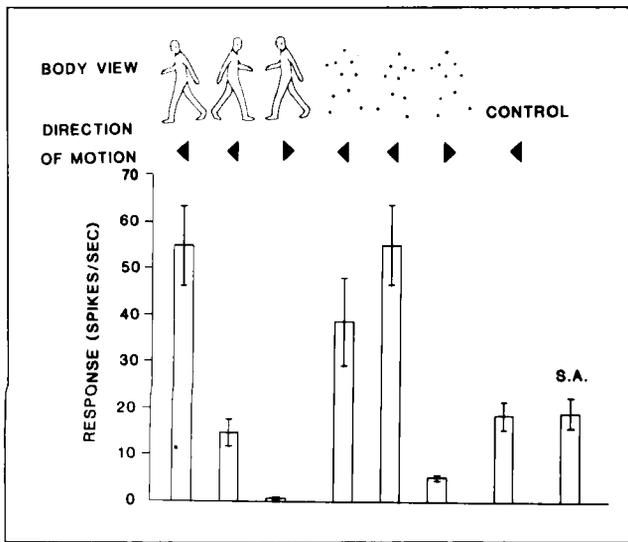


Figure 1. Selectivity for direction of body movement but not body form for biological motion stimuli. The mean response \pm SEM of one cell (D222_22.45) to walking bodies under normal and biological motion conditions. The upper section shows schematic representations of the view and direction of movement of stimuli. Under normal lighting, the cell responded to the sight of the left profile body view walking (compatibly) to the monkey's left. The sight of the right profile view walking (incompatibly) to the left gave no response above the cell's spontaneous activity (SA) ($p > 0.05$) and the optimal body view moving to the right produced inhibition relative to SA ($p = 0.03$). Under biological motion conditions both left and right body views moving in the cell's preferred direction elicited a response greater than control movement and SA ($p < 0.02$ each comparison). Left and right directions of motion were discriminated under biological motion conditions. Overall effect of conditions, $F(7,34) = 14.0, p < 0.0005$; number of trials for each condition, left to right, $n = 5, 3, 5, 7, 10, 4, 5, 3$.

Two of the four gave responses to moving dot stimuli that were statistically indistinguishable from responses to stick figure stimuli; the remaining two cells responding only to stick figure stimuli.

Most of these cells (14/18, 78%) that showed statistical discrimination between directions and body views also showed a reduction in absolute response magnitude relative to the walking stimuli under natural lighting. Cells with this type of response characteristic were found for both moving dot and stick figure stimuli. Figure 2 shows an example of this type of response to stick figures. The cell was selective specifically for the front view of the body walking away from the monkey (incompatible movement). With stick figures, a reduced response was found (latency approximately 100 msec) to the preferred stimulus but it was still significantly above spontaneous activity and controls (not shown). Stick figure equivalents of noneffective walking stimuli (e.g., the back view of the body moving away from the monkey) produced significantly smaller responses.

Four cells (6% of the total sample of 72 cells, 22% of the cells responding to biological motion stimuli) responded to the biological motion stimuli in a manner

that was very similar to the responses to the real walking stimuli. Figure 3 shows the responses of one cell to real and stick walking figures. As can be seen, the cell has a preferred stimulus of compatible walking to the monkey's right. The left profile walking in the preferred direction and the preferred body view walking to the right both produced significantly weaker responses. The stick figure responses also followed this pattern, with no significant differences found across comparable body view/direction of movement combinations between real and stick figures. As can be seen, the response latency under both biological motion conditions and normal lighting is approximately 100 msec. Figure 4 shows the responses of another cell to real and dot figure stimuli. This cell was selective for the left profile view moving to the left. As can be seen this selectivity was maintained at comparable levels when biological motion stimuli were used.

Jumbled Articulation

As an additional investigation of form selectivity, comparison was made of responses to natural and jumbled configurations of the biological motion stimuli. The jumbled figure stimuli (see Methods) contain the same rigid linkage structure as the biological motion stimuli, the same overall translation vector, and the same component vector of each point. They differed only in the relative positions of the light points. A total of 14 cells were tested with these randomized moving dot displays (jumbled figure). Of these, 10 cells proved to be insensitive to biological motion stimuli and the jumbled figure. For the three cells where a selective response was seen to the biological motion stimuli (i.e., preferentially responding to one body view and direction combination displayed in biological motion conditions), the response to the jumbled figure moving in the preferred direction was significantly ($p < 0.05$) reduced compared with the preferred view and direction combination. The one cell that was selective for direction but not body view under biological motion conditions also responded to the jumbled figure moving in the preferred direction. Figure 5 shows the response of a cell to biological motion stimuli and a jumbled biological motion stimulus. As can be seen, the response differentiates the preferred movement (left profile walking to the right) from compatible movements to the right and the left profile walking left. The response to the biological motion stimuli was reduced compared with the comparable real stimuli (approximately 50%). The jumbled articulation stimulus produced a response that was no greater ($p > 0.5$) than the cell's spontaneous activity. Thus this cell shows statistically reliable discrimination between not only two alternative body view representations, but also between the preferred view and direction combination (as a biological motion stimulus) and the jumbled figure equivalent.

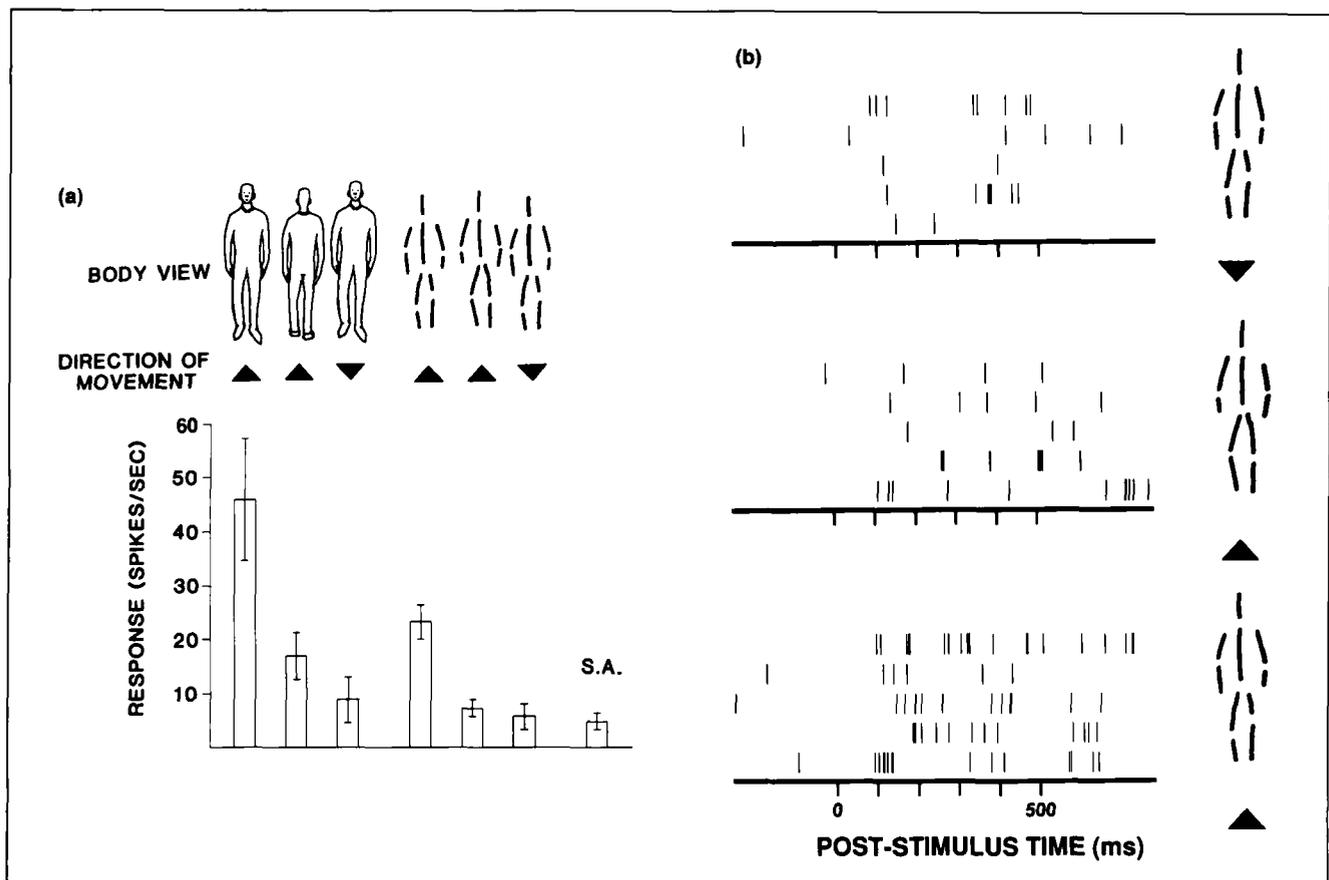


Figure 2. Reduced response to stick figures. **(a)** The mean response \pm SEM of one cell (J039_28.14) to walking bodies under normal and biological motion (stick figure) conditions. The upper section shows schematic representations of the view and direction of movement of stimuli. Upward arrows indicate movement away from the monkey, downward arrows indicate movement towards the monkey. The view and direction of the stick figures maintain the same order as the normal stimuli. Under normal lighting the cell responded more to the front view walking away from the monkey than to other view/direction combinations, including back view approaching (not shown) ($p < 0.05$ each comparison). The responses to the stick figure walking stimuli gave statistical discrimination between view and direction combinations ($p < 0.05$) that mirrored the responses to the real stimuli but were reduced. SA = spontaneous activity. Two-way ANOVA, overall effects: testing conditions (normal vs. stick figure) $F(1,52) = 5.53, p < 0.025$; view/direction (front/rear, approach/retreat) $F(3,52) = 8.56, p < 0.005$; interaction $F(3,52) = 0.673, p = 0.57$. Number of trials per condition $n = 10$ for normal stimuli, $n = 5$ for stick figures. **(b)** Rastergram displays of the five trials for each of the biological motion stimulus conditions. Each trial is represented by a single row of ticks, each tick indicates one action potential. Poststimulus time is given at the figure base. Note the latency of approximately 100 msec.

Eye Movements

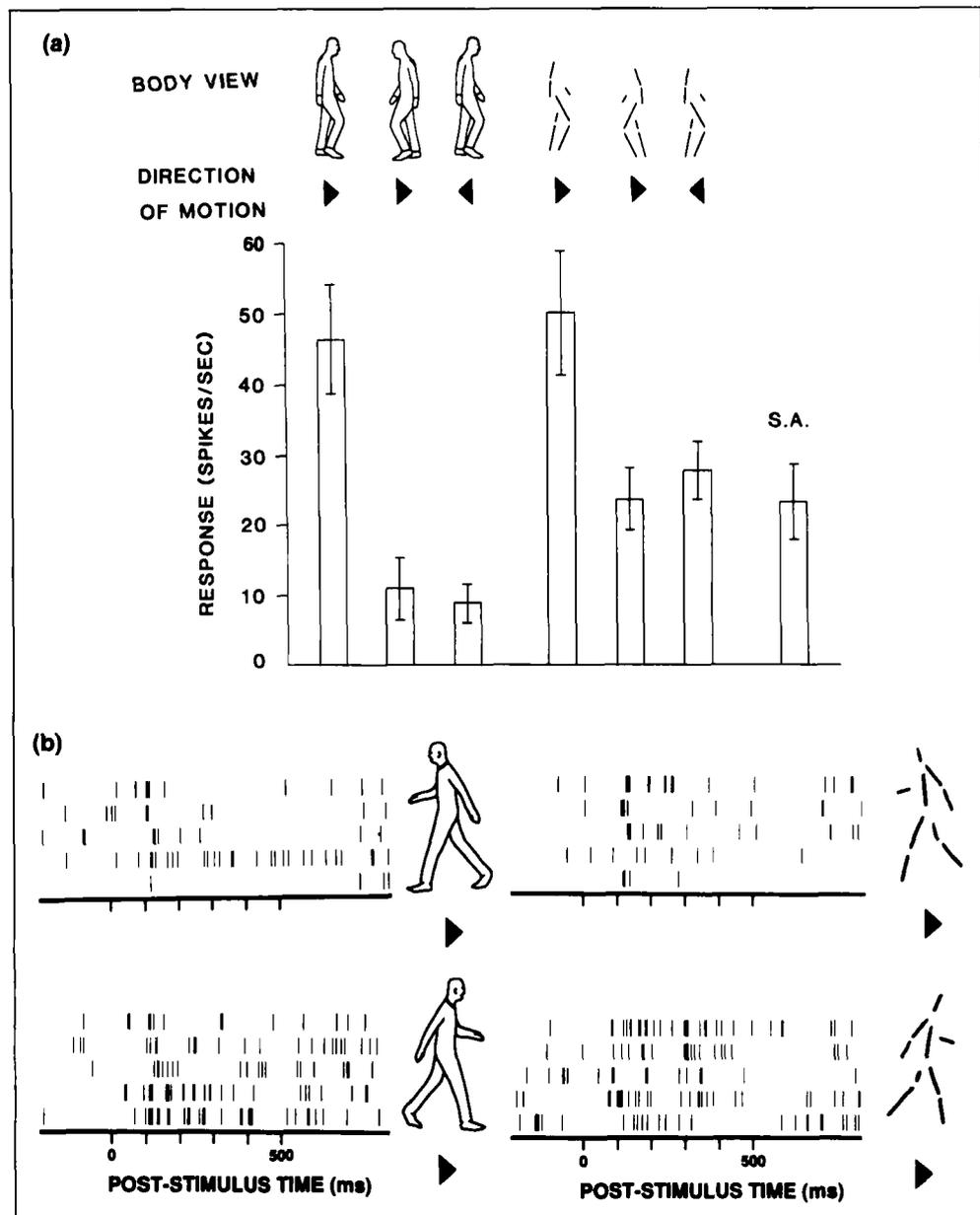
Figure 6 shows a typical example of the eye movement recordings and the response to individual trials for one cell. The upper figure shows the responses to an effective stimulus (for this cell biological motion walking compatibly to the monkey's left). As can be seen from the eye movements just prior to the stimulus onset (time 0), the monkey saccades to the LED and maintains fixation until approximately 250 msec poststimulus onset. For each of the five presentations, the cell response occurs with a latency of approximately 150 msec. The lower figure shows the eye movements and cell responses to the five trials of the incompatible movement (i.e., walking backward to the monkey's left). Again, the eye movements show maintained fixation but there is clearly no cell response in any trial. Indeed there is some evidence for inhibition to this type of stimulus. In both the compatible and incompatible stimulus conditions, there is

evidence for smooth pursuit eye movement (following the "wrist dot" down) for both stimuli after the initial fixation period. Since the eye movements are comparable for both stimuli, they cannot account for the difference in response magnitudes: the only difference is in the presented stimulus. Furthermore, for the effective stimulus it can be seen that despite small variations in eye position the response onset is tightly time locked (Fig. 6, upper). A similar lack of relationship between eye movements and response selectivity was obtained for cells from all four recording subjects.

Discrimination Measures

In order to examine the discrimination shown by the responses of the tested cells to the differing stimuli, we calculated discrimination measures for both direction

Figure 3. Responsiveness to stick figure stimuli. **(a)** The mean response \pm SEM ($n = 10, 10, 10, 5, 10, 10, 10$) of one cell (J067_2794) to walking bodies under normal and biological motion (stick figure) conditions. The upper section shows schematic representations of the view and direction of movement of stimuli. The cell responded to the right profile walking to the right, both for normal and stick figure stimuli. Responses to inappropriate views or directions were not significantly different from the spontaneous activity (SA) or control stimuli (not shown). Two-way ANOVA showed a main effect for four tested view/direction combinations [$F(3,62) = 11.3, p < 0.0005$] but not for lighting condition (normal vs. sticks) [$F(1,62) = 1.1, p = 0.29$]. The interaction was nonsignificant [$F(3,62) = 2.5, p = 0.07$]. **(b)** Rastergram display showing five responses to normal lighting condition stimuli (left) and the responses to the biological motion equivalents (right).



(I_d) and view (I_v). These were calculated as $I_d = 1 - (R_{oppd}/R_{pref})$ and $I_v = 1 - (R_{oppv}/R_{pref})$, where R_{pref} = response to the preferred view and direction combination - spontaneous activity (SA), R_{oppd} = response to the preferred view moving in the opposite direction from the preferred direction - SA and R_{oppv} = the response to the view opposite to the preferred view moving in the preferred direction (- SA). The preferred direction and view were first defined under normal lighting and then the magnitudes of the responses R_{pref} , R_{oppv} , and R_{oppd} measured and compared under biological motion. The distribution of I_d values is shown in Figure 7a and the distribution of I_v values is shown in Figure 7b. The upper sections show the values obtained under normal lighting conditions; the lower section shows the values calculated for the same cells using biological motion stimuli. The black bars indicate

the direction (Fig. 7a) and view (Fig. 7b) indices for those cells which showed statistical discrimination for direction and view under biological motion. Note that it would be expected that cells lacking selectivity under biological motion would have a wide range of discrimination index values because the cell responses would fluctuate around spontaneous activity levels.

Comparison of the direction indices for all cells revealed a reduction in I_d for biological motion stimuli compared with I_d for natural stimuli (medians = 0.71 and 0.84, respectively, Wilcoxon test, $W = 196, N = 36, p = 0.03$). A more marked reduction was found for I_v under biological motion compared to normal lighting (medians = 0.20 and 0.68, respectively, Wilcoxon test, $W = 89, N = 43, p < 0.0005$). (Values were included only for cells where the responses differed from spontaneous activity by more than 1 spike/sec.) The same

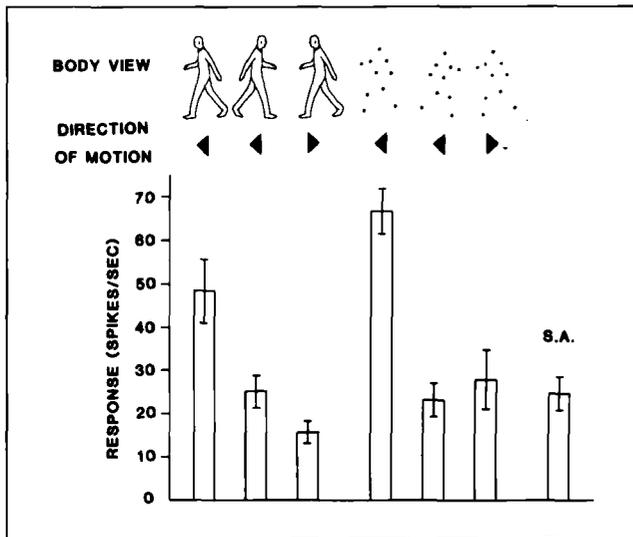


Figure 4. Responsiveness to biological motion. The mean responses \pm SEM are shown to the stimuli depicted above (cell J060_2989). The cell's selectivity for compatible walking to the left is maintained with biological motion stimuli. Two-way ANOVA showed a main effect for view/direction combination [$F(3,32) = 14.0, p < 0.0005$] but not lighting conditions (natural vs. biological motion dots) [$F(3,32) = 3.1, p = 0.09$]. The interaction was nonsignificant [$F(3,32) = 1.16, p = 0.34$].

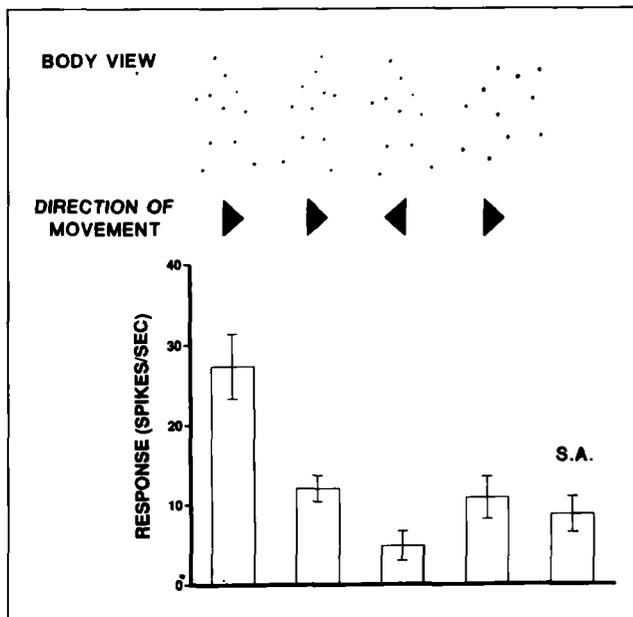


Figure 5. Discrimination between normal and jumbled articulation (cell J053_23.55). The upper section depicts the stimuli. Under biological motion (and natural) conditions the cell responded selectively to the left profile view of a body walking backward. Response to the jumbled biological motion stimulus moving to the right was no different from the response to the moving point light display of the right body view walking to the right ($p > 0.5$) but less than the biological motion representation of the preferred stimulus under natural conditions ($p = 0.003$). [One-way ANOVA, overall effect of conditions: $F(4,50) = 8.15, p < 0.0005$.]

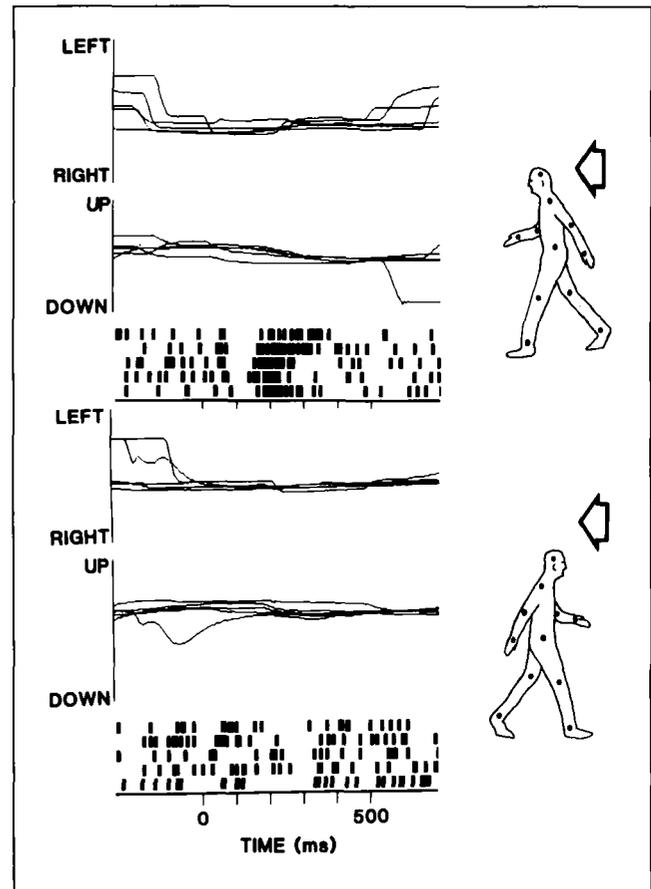


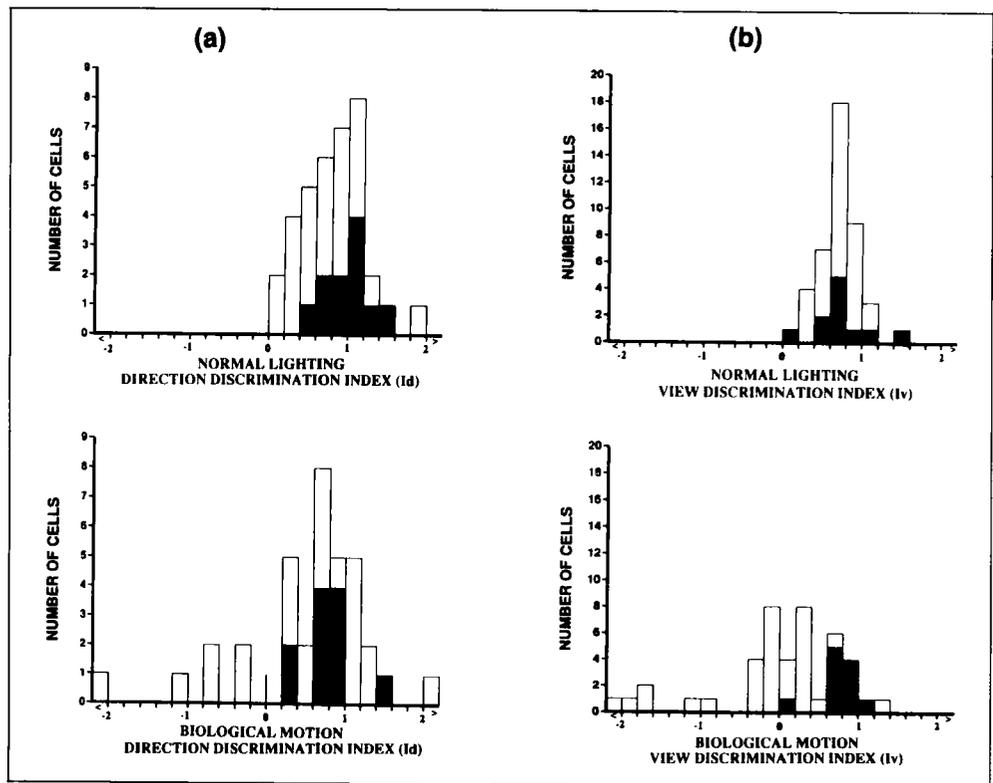
Figure 6. Eye movements do not account for selectivity of biological motion. The cell responded to compatible walking to the left. Responses to individual trials of biological motion stimuli are shown in the lower sections, while both horizontal and vertical eye movements are shown for each trial in the upper sections. Full scale deflection = 50° for horizontal and vertical traces (eye movements were recorded over a range of positions $\pm 20^\circ$ from straight ahead; clipping occurred outside this range). The rastergrams show a good response to each of the trials with compatible walking (left profile walking left) but no response to any of the incompatible walking trials (right profile body view walking left). For both stimulus conditions the eye movements were comparable and thus cannot account for the large differences in response magnitudes.

comparisons of the indices were also performed for only those cells whose responses under biological motion stimuli showed statistical discrimination for both view and direction. A small drop was observed in I_d [medians = 1.00 (normal lighting) and 0.79 (biological motion), Wilcoxon test, $W = 9, N = 11, p = 0.04$]. Surprisingly, the view discrimination index I_v under the two conditions showed no significant difference (median for natural lighting = 0.69, biological motion stimuli = 0.80, Wilcoxon test, $W = 27, N = 11, p > 0.5$).

Cells Selective to Whole Body Movements Other Than Walking

Sensitivity to biological motion stimuli was found not only for cells selective for ambulation but also for cells

Figure 7. Distribution of view and direction discrimination indices. **(a)** Distribution of the direction discrimination (I_d) index for all cells selective for body view and direction and tested under natural lighting (upper) and biological motion (lower) conditions. [$I_d = 1 - (R_{\text{oppd}}/R_{\text{pret}})$], where R_{pret} = response to preferred view and direction - spontaneous activity, R_{oppd} = response to the preferred view moving in the opposite direction - spontaneous activity]. **(b)** Distribution of the view discrimination measure (I_v) for cells tested under natural lighting (upper) and biological motion (lower) conditions. [$I_v = 1 - (R_{\text{oppv}}/R_{\text{pret}})$], where R_{pret} = response to preferred view and direction - spontaneous activity, R_{oppv} = response to opposite view moving in the preferred direction - spontaneous activity.] The black bars indicate cells that showed statistical discrimination between view and direction under biological motion.



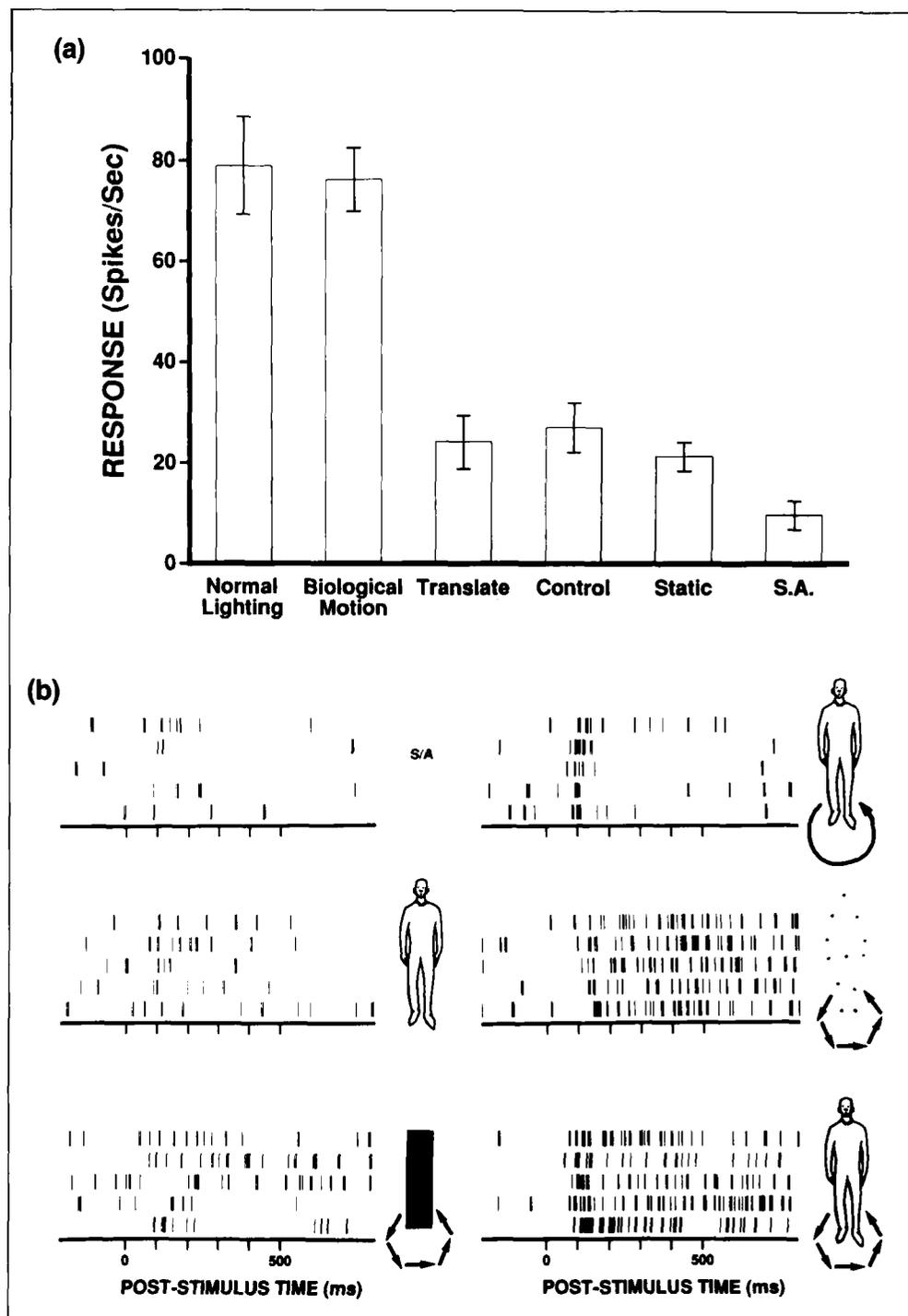
during other whole body actions. Ten cells were tested with biological motion stimuli that were responsive to other whole body movements. Five of these were completely unresponsive to biological motion stimuli, whereas the other five showed either reduced or comparable response magnitudes and response patterns to those obtained under normal lighting conditions. Figure 8 shows an example of a cell that was selective for the sight of the whole body rotating. While the direction of rotation did not matter (not shown), controls of a comparable size rotating at similar speeds with component articulation did not elicit a response. Biological motion stimuli produced responses comparable with the live rotations. This cell was of interest because, when tested with the rotation without articulation (achieved by having the experimenter standing rigid on a rotating platform under normal lighting), the cell showed only a very weak response. This implies, that for this cell, the sight of limb articulation during whole body rotation was necessary. It also was apparent that biological motion conditions were sufficient to elicit a strong response. The rastergram display (Fig. 8b) suggests that the response consists of two components. The first of these is transient (lasting 10–40 msec) and can be elicited by the sight of rigid body rotation. However, the following sustained response can be elicited only by nonrigid rotation of the body.

A second example of a cell selective to whole body movement downward defined by the pattern of articulation which was also sensitive to biological motion stimuli is shown in Figure 9. Translation of a nonarticulating body (a life-sized 2-D model) down produced a nonsignificant response that was comparable to control movement down. Thus for this cell the articulation was a necessary component of the stimulus. The cell's response to the biological motion version of the body moving down (with articulation) was similar to the response to the same stimulus under normal lighting. Thus for this cell it was shown that the relative movements between the points of articulation produced by crouching down were necessary and sufficient to produce the maximal cell response.

Location of Cells

Figure 10 shows the histological reconstruction of the positions of the tested cells in the central region of the recording area in one subject (monkey J). The left columns show the locations of all cells tested for sensitivity to biological motion stimuli. The right columns show only those cells that responded selectively to biological motion (either dots or stripes or both). As can be seen

Figure 8. Response selectivity to biological motion stimuli for rotating bodies. **(a)** (Cell J120_28.22) The cell responded to nonrigid whole body rotation under normal and biological motion conditions. Responses were significantly lower to control objects (matched for size) rotating with the same direction and speed, the static body view and a rigid body rotation ($p < 0.0005$ each comparison). There was no significant difference between biological body rotation under biological motion and normal lighting conditions ($p = 0.73$). Overall effect of conditions: $F(5,24) = 27.1, p < 0.0005$. **(b)** Raster-gram display of the cells' responses to the conditions shown in **(a)**. Top left: presentation of the LED alone (S/A or "no stimulus" condition). Mid left: static body. Bottom left: control rotating nonrigidly. Top right: rigid rotation of the body. Mid right: biological motion display of articulating body rotation. Bottom right: body rotating nonrigidly under normal lighting. The broken arrows indicate nonrigid rotation whereas the smooth arrow indicates rigid body rotation.



in the right hemisphere, cells selectively responding to biological motion stimuli were found in both the upper bank and the fundus of the sulcus. In the left hemisphere, the cells tested for sensitivity to biological motion were located only in the upper bank of STPa (areas TPO and PGa of Seltzer & Pandya, 1978). We have no reason to assume that the distribution of sensitivity to biological motion stimuli showed any hemispheric differences. Reconstruction of the other subjects indicated that cells sensitive to biological motion were located in the same regions.

DISCUSSION

Summary of the Results and Mechanisms of Cellular Sensitivity to Biological Motion Stimuli

One-third of the cells (25/72) selective for the form and motion of walking bodies that were tested showed partial or full sensitivity to biological motion (either dot or stripe) stimuli. Of the cells that did respond, one-third (7/25) showed only partial sensitivity to form from motion, in so far as the responses under biological motion

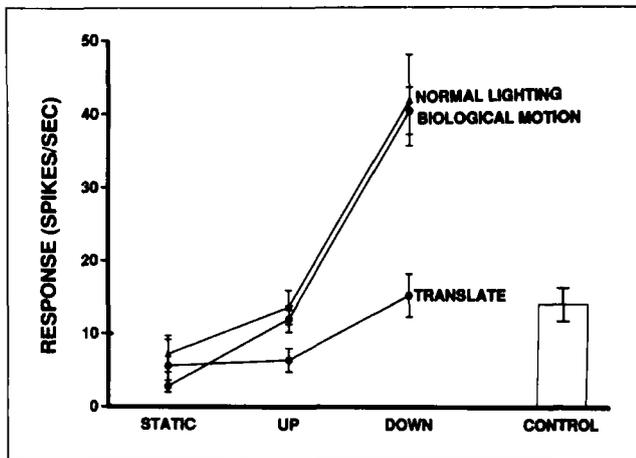


Figure 9. Articulation as a necessary and sufficient condition for response to body motion (cell J045_26.46). Crouching down, either as a biological motion stimulus or under normal lighting, produced a response significantly greater than reverse direction of body motion (standing up), static body views, and control movement downward ($p < 0.05$ each comparison). [Two-way ANOVA main effect of motion type (down/up/static) $F(2,46) = 45.0, p < 0.0005$; main effect of stimulus type (translate/natural/biological motion) $F(2,46) = 7.89, p = 0.001$; Interaction $F(4,46) = 2.66, p = 0.045$]. Downward translation of the head and body (without articulation) did not produce a response different from control motion downwards [$t(8) = 0.32, p > 0.75$].

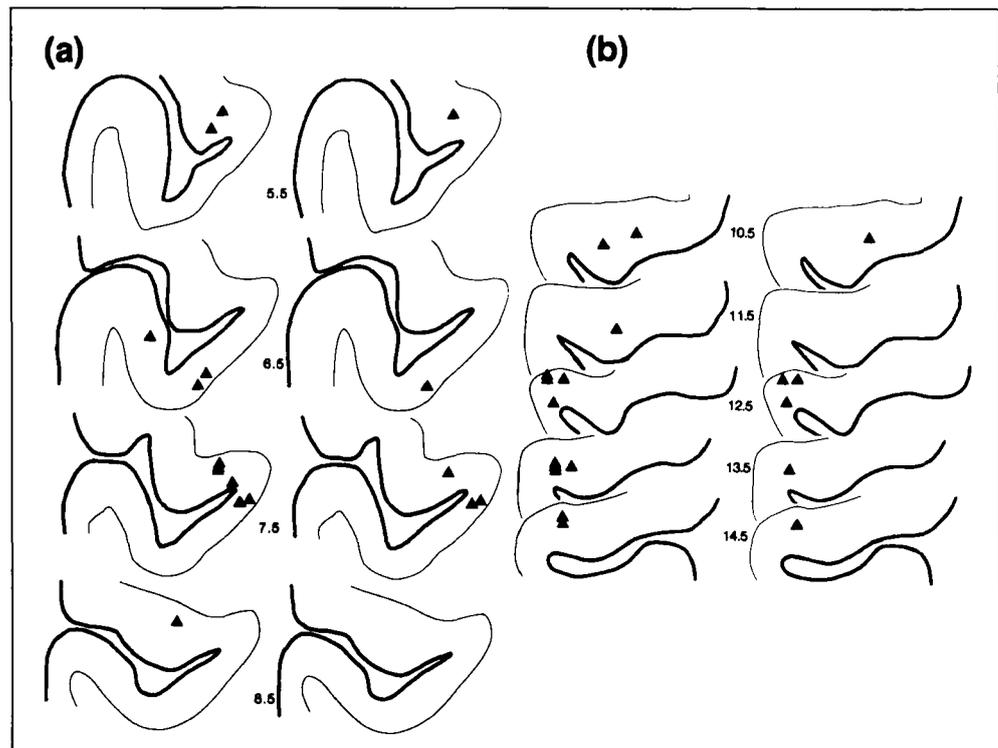
conditions maintained direction sensitivity but failed to discriminate body view. For these cells a limited capacity to process "body form" was indicated by the observation that often responses to all views of the walking body

depicted in biological motion were greater than moving dot or stripe control stimuli.

The majority (18/25) of cells responding to biological motion stimuli showed statistical discrimination for both direction of movement and body view. Such cells typically showed reduced responses to the impoverished stimuli compared to full light conditions. Response reduction is perhaps not surprising given the loss of contour information in the biological motion stimuli. Some (4) cells showed response magnitudes and selectivity to biological motion stimuli that were statistically indistinguishable from those to the "real" stimuli under normal lighting. All 18 cells were thus sensitive to detailed form information (body view) from the pattern of articulating motion present in biological motion stimuli. The cell responses thus provide direct evidence for neural mechanisms computing body form from nonrigid motion.

In relation to the possible schemes of processing described in the introduction, the data here presents evidence in favor of scheme (2) whereby form is calculated from motion inputs alone (in particular see Figs. 8 and 9). Although all three mechanisms outlined in the introduction could contribute redundant information to the ultimate perception of body motion, under biological motion conditions only mechanism (2) would contribute to the perceptual ability to differentiate body form [schemes (1) and (3) would predict only responses equivalent to nonrigid controls moving in the preferred direction]. As noted in the introduction, area STPa receives inputs from both areas MST and FST (Boussaoud et al., 1990). However, the number of cells (6%) showing comparable responses to natural lighting and biological

Figure 10. Histological reconstruction of recorded cells. (a) Series of four sections (at 1 mm intervals) of the superior temporal sulcus (STS) from the right hemisphere of one monkey (J). (b) Series of five sections showing only the upper bank of the STS from the left hemisphere of the same monkey. The thick line indicates the surface of the brain, the thin line marks the boundary between gray and white matter. The left columns of a and b show the location of the cells tested for sensitivity to biological motion. The column on the right shows the location of those cells that showed statistically significant discrimination for the form of biological motion stimuli. The figures between the two columns indicate the distance in millimeters anterior to the interaural plane.



motion stimuli is small: the majority of cells responding to the impoverished biological motion stimuli showed reduced response magnitudes. This, and the observation that nearly all cells responded to translation of the appropriate body form in the preferred direction (Oram et al., in prep.) led us to propose a tentative model for the derivation of cellular selectivity to biological motion stimuli (Oram & Perrett, 1994). In particular, it is suggested that scheme (1) of the Introduction is prevalent (Oram et al., in prep.), with cell selectivity to form and motion (body walking) resulting from integration of form inputs and motion inputs from separate dorsal and ventral sources. For a minority of cells it is suggested that the motion inputs also include local field (possibly from MSTl) as well as wide field inputs (from MSTd). We propose that it is the learned association of local field inputs (potentially coding relative motion of light points) with the overall translation (wide field motion) and form inputs that gives rise to sensitivity to biological motion stimuli. While this tentative model would explain many of the results reported here, we stress that the cell selectivity to biological motion stimuli reflects the ability of cells in the macaque STPa (after "learning") to compute form from the motion inputs alone (scheme 2 of the Introduction), and does not rely on the presence of form inputs.

Sensitivity to Global Motion Patterns

Populations of cells in the anterior sections of the temporal lobe have been found sensitive to the movement of individual limbs (Perrett et al., 1985b, 1989a,b, 1990b; Hasselmo et al., 1989). The cells reported here, however, responded only to whole body motion and not single limb articulation. It is unlikely therefore that the sensitivity observed to biological motion stimuli can be accounted for in terms of isolated patterns of local relative motion. The global nature of the motion analysis was indicated by the discrimination of body view for whole body movements in the same direction and by the observations that cells were (1) unresponsive to control patterns of dots moving nonrigidly and (2) could respond differentially to jumbled and normal biological motion stimuli. These observations indicate the complexity of the analysis being performed, since all connected pairwise relative motions of individual limbs remain in the jumbled and opposite view stimuli, yet cells did not respond. An analogous situation exists with some cells selective for static views of the head. These cells discriminate between different views with the same facial features (e.g., left and right profile) and they also respond less to the presentation of a jumbled face even when all the facial features are present (Perrett et al., 1982, 1991, 1992; Perrett, Mistlin, Chitty, Smith, Potter, Broennimann, & Harries, 1988).

Relationship of Eye Movements and Cell Responses

It could be argued that the cell selectivity we observed was related to eye movements or position. However, we believe this to be extremely unlikely for the reasons given below.

1. The STP area is a large structure and extends from the parietal lobe to the temporal pole and it is divided into at least two functionally distinct subregions (STPp and STPa). Our recordings were restricted to STPa. One might expect eye movements to have some influence in STPp (given its proximity to visuomotor areas within parietal cortex and its inputs from MST). Indeed there have been brief reports that some STP cells show differential responses dependent on eye movements (Colby & Miller, 1986; see Colby, 1991 for one example). Their studies indicated only 6% (5/90) of cell responses were found to be related exclusively to eye movements (10% were visually responsive as well as sensitive to eye movements, 4% responded to the visual stimuli but ceased firing when eye movements were made, 80% showed no modification of response with eye movements; C. L. Colby, personal communication). Thus the proportion of eye movement related activity decreases markedly from MST to STP. This decrease is likely to continue along the temporal sulcus from the parietal lobe toward the temporal pole. It is not clear if the cells Colby and colleagues recorded were located in the posterior (STPp) or anterior (STPa) sections of STP. Therefore the figure of 6% provides an upper limit to the proportion of cells in our study whose response selectivity might be accounted for by differential eye movements. We found 37% of the cells studied responded differentially to biological motion stimuli. This figure is far greater than the 6% found by Colby. Even in the absence of eye position recording, the selectivity for biological motion stimuli is unlikely to be attributable to eye movements.

2. The receptive field size of cells in STPa is very large and typically covers the fovea (Bruce et al., 1981). Similar selectivity for static stimuli at different positions within the large receptive fields has been reported for cells in inferotemporal cortex and STPa (Desimone et al., 1984; Gross, 1992; Tovee & Rolls, 1993). Although we did not expressly check the receptive fields for all cells, of those cells which receptive fields were mapped, we have observed similar positional invariance within the STPa. Selectivity for static and moving stimuli was maintained to eccentricities of 10–20° either side of the fovea (Perrett et al., 1989b; unpublished studies Perrett, Harries, & Oram). Thus with large receptive fields and positional invariance, difference in eye positions ($\pm 20^\circ$) is unlikely to have effected response selectivity in this study.

3. From Figure 6 it is evident that despite small variations from trial to trial in eye position ($\pm 3^\circ$), the effective biological motion stimulus always produced a clear response with very similar latencies. Furthermore when

the monkey was looking within the same range of positions there was never a response to the ineffective stimulus. Given the vertical and horizontal extent of the stimuli (approximately 10 by 5°) and the likely receptive field size (at least $\pm 15^\circ$), both effective and ineffective test stimuli would have fallen within the same range of positions and well inside the cell's receptive field.

4. We have examined all eye position traces of cells during testing of biological motion stimuli and their derivative (velocity). There was evidence of tracking for some recordings, e.g., Figure 6, however, the velocity range of eye motion was found not to differ across the stimulus conditions. We therefore can see no explanation for the response selectivity other than the difference between the effective and noneffective stimuli (e.g., left vs. right body profile in Fig. 6).

In summary these arguments indicate that it is unlikely that the selectivity for biological motion stimuli that we observed was due to eye movements. First, the available evidence suggests that cell responses in STPa are generally unrelated to eye movements. Second, given the size of STPa cell receptive fields and positional invariance, any small variation in eye position would not account for differential responses. Third, direct measurements of eye position indicated that differences in eye position/velocity across stimulus conditions were indeed small. Finally, and more importantly, there was no consistent relation between eye position/velocity and neural responses reported here.

Motion Processing in the Ventral Visual Areas

The stream of visual processing running ventrally into the temporal cortex is commonly thought to be associated with the encoding of object form. The specification of an object's form is usually thought to involve an analysis of static visual information. Indeed lesion studies have indicated that temporal cortex is needed for the learning and memory of static patterns (Dean, 1976; Ungerleider & Mishkin, 1982). Processing of static form in this region is also indicated by the finding of single cells which exhibit a high degree of selectivity for static objects (see Introduction).

We have shown that neural sensitivity to form and motion does not depend solely on form visible at any particular instant but can be generated from motion information alone [scheme (2) of the Introduction]. The computation of form from motion may well involve or depend on processing conducted in the dorsal stream of processing. Certainly lesions to the dorsal system (MT/MST) can produce impairment in the extraction of shape from motion (Andersen & Siegel, 1989; Siegel & Andersen, 1986). The properties studied here could well depend upon the projections from the motion processing

areas (MT/MST/FST) into the cortex of the STS (Felleman & Van Essen, 1991).

Lesions of the inferior temporal cortex in monkey impair the ability to learn shape discrimination where shape is defined by the relative translation of random dot patterns (Britten, Newsom, & Saunders, 1992). Again this finding indicates the utilization of movement information to define form within the ventral stream, though this particular processing capacity could depend on contour analysis performed in area V4.

Relation to Neuropsychological Studies

It is becoming increasingly apparent from neuropsychological studies that recognition using static and dynamic visual cues is dissociable. Impairments in the ability to recognize facial expression from static photographs (Ekman & Friesen, 1976) do not necessarily parallel recognition impairments for expression displayed in biological motion format with light dots attached to the face (Bassili, 1979; Humphreys, Donnelly, & Ridloch, 1993; for discussion see also Campbell et al., 1992).

Neuropsychological studies also indicate that human brain mechanisms involved in the processing of complex motion (such as body form defined by biological motion and the form of cylinders defined by rigid rotation) can be dissociated from mechanisms involved in processing of direction and velocity (Vaina, Lemay, Bienfang, Choi, & Nakayama, 1990). More dorsal lesions are associated with a loss of simple motion processing, whereas lesions more anterior and ventral are associated with disruption of form from motion. A further example of this dissociation is provided by Patient LM (Zihl, Von Cramon, & Mai, 1983) who has been described as "motion blind" following lesions to dorsal visual areas. LM cannot track movements at velocities greater than 8°/sec; fast moving objects appear to her as a series of static images. Despite this dramatic motion processing deficit LM retains some capacity to recognize body form defined by biological motion stimuli (McLeod, Zihl, Perrett, & Benson, unpublished studies, 1990).

Relationship to Computational Models

Ullman's algorithm for extracting form from motion could apply to biological motion stimuli except that it requires four visible noncoplanar points on each rigid element (Ullman, 1979). The earliest computational model to calculate an object's linkage structure from biological motion displays (Rashid, 1980) used the correlation of position and velocity of dots in successive video frames to postulate the rigid connecting links between the dots. This simple procedure produced reasonable solutions for simple stimuli (an idealized walking man). For complex stimuli (e.g., two men walking around one another) the procedure was slow and inaccurate. More recent computational approaches (Hoffman

& Flinchbaugh, 1982; Bennett & Hoffman, 1985; Sugie & Kato, 1987; Webb & Aggarwal, 1982) make use of natural constraints which are likely to exist in the stimuli. For example Webb and Aggarwal (1982) assume the axis of rotation of each locally rigid element remains fixed during the rotation. The resolved trajectory for one rod element (e.g., the torso) can be used as a frame of reference for defining the trajectory of the next linked rod element (upper arms and leg sections). Such approaches can resolve the correct linkage in biological stimuli extremely efficiently, indeed performance can reach the theoretical limit of three successive frames providing no assumptions are broken (e.g., Sugie & Kato, 1987).

It is relevant to consider the data from psychophysical studies that indicate that naive human observers may perform less efficiently than the recent computational models. Naive observers can correctly identify a biological motion stimulus with exposure durations of between 0.1 and 0.2 sec (4–8 frames, Johansson, 1976; Lappin, Doner, & Kottas, 1980). With computer-animated biological motion displays subjects can discriminate normal walking figures from jumbled figures where the position of limb marker points has been moved randomly a distance 30% of the head to ankle height (Perrett et al., 1990a). Observer performance on such discrimination tasks is profoundly affected by the presence and type movement of background masking dots (Cutting et al., 1988; Perrett, et al., 1990a; Proffitt, Bertenthal, & Roberts, 1984), unlike the computational models that should have no problem with masking dots. Naive subjects perform the normal/jumble discrimination task initially rather poorly and often require more than 8 frames to perceive the figures. Minimal practice (30 trials) substantially improves performance. Even in the presence of background masking dots, which remove residual static form cues, experienced subjects can perform above chance with 2 to 3 frames exposure.

We learn from these perceptual studies in humans that purely dynamic cues can be used to retrieve structure extremely quickly. Considering STPa cell response latencies similar conclusions can be reached. Although detailed studies of the response time course have yet to be made, it is apparent that cell responses to biological motion stimuli can occur within 150 msec after stimulus onset (Fig. 6).

Computational models derived so far for interpretation of biological motion have two properties that make them inadequate for accounting for psychophysical and single cell data. The first property is that the models are general purpose. The perceptual system, however, appears to employ specific mechanisms rather than a general purpose analysis. Sumi (1984) and Dittrich (1993) found that normally oriented biological motion stimuli were more accurately perceived than inverted stimuli. If the visual system employs a general purpose analysis then perception should be equally successful in identifying inverted

or upright figures. The physiological data indicate a much higher degree of specialization. Even for the normally experienced upright orientation, different cell populations are employed for the analysis of different types of body motion (walking, crouching, rotating). At a more detailed level, for each type of movement (e.g., walking) subpopulations of cells are involved in the analysis of specific directions of movement and specific body views (e.g., left profile view walking left). In all, eight subpopulations would be needed to cover bipedal walking along the horizontal plane (two types of walking, forward and backward, in four directions, left/right and toward/away). Thus while the majority of computational models apply equally to all perspective views, the brain systems involved in computing biological motion appear to employ view and direction specific neural mechanisms.

The second important difference between neural mechanisms studied here and the computational approaches is that many of the models achieve a less complete description of the visual input compared to natural recognition systems. Many computational schemes retrieve only the linkage structure (which element connects with which) whereas the systems studied here are capable of providing additionally information about the nature of the linked stimuli. The cell responses can, for example, provide evidence that the stimulus is a body (as opposed to other objects or a jumbled body), that it is walking (not rotating or crouching), and more specifically that it is seen from left profile view and is walking to the left.

The improvement of human perceptual performance with practice indicates that the processing of biological motion stimuli may in some way involve “top-down” influences where expectations for the form of the moving object are compared against visual input. The appropriate computational model for processing would appear to be one in which input data are checked against specific models stored in memory and the results of the matching used to guide subsequent predications (see Lee & Chen, 1985; Leung & Yang, 1987). A role for top-down influences has also been suggested for object recognition (Lowe, 1987; Seibert & Waxman, 1991, 1992a,b). It remains to be determined what role experience has in shaping STPa cell responses to biological motion stimuli (see above).

Hybrid computational models might be more appropriate for describing the cellular responses to biological motion stimuli. Such models could perhaps first search for potential links in the articulating array and then check these against specific stored representations of the static or articulating bodies (see Lee & Chen, 1985; Leung & Yang, 1987). The stored representations could be object centered (Marr, 1982; Marr & Vaina, 1982; Marr & Nishihara, 1978; Lowe, 1987) or, more in agreement with the physiological data, viewpoint dependent (Koenderink & van Doorn, 1979; Seibert & Waxman, 1991; Goddard, 1992).

METHODS

Four subjects were used (*Macaca mulatta*, 3 male B, D, H, weight 5–8 kg, 1 female J, weight 4 kg from a U.K. registered breeding colony). The subjects were trained to fixate on LED attached to a plain white wall 4 m away. The behavioral response was to lick for fruit juice to a green light and refrain from licking to a red LED to avoid a weak saline solution. During this discrimination task, the subject was placed in a primate chair. A half second warning tone was given before each trial, then the LED was turned on. The color of the LED was varied in pseudo-random order across trials under computer control. Videodisc sequences or real 3-D moving objects were presented either to cross the LED or projected to cover the LED at each trial.

After surgery under pentobarbital (Sagatal) anesthesia (with full sterile precautions) to implant a recording chamber (see Perrett et al., 1985a for details), the monkey was allowed to recover, retrained on the LED task until performance was greater than 80%, then experiments were started. Standard chronic recording techniques were used to record from single cells in the STPa (areas TPO and PGa of Seltzer & Pandya, 1978) when stimuli were presented. Spikes from individual cells were discriminated using a threshold voltage window. The threshold was set manually for each cell tested. Spike data were stored in 5 msec time bins. Responses were measured as spike frequency estimated from the period 100–350 msec poststimulus onset.

Eye movements were recorded throughout the stimuli presentations using an infrared corneal reflection system (ACS, modified to allow both vertical and horizontal position to be recorded from one eye). The analogue output was sampled at the same frequency as the spike signals with 8 bit accuracy over the range $\pm 20^\circ$ and stored with the spike data for each trial.

Stimuli

The stimuli were either real 3-D presentations or sequences of frames on a videodisc. They included images of the experimenter walking, both forward (compatible movement) and backward (incompatible movement) in different directions (toward and away from the monkey and moving to the monkey's left and right). The biological motion stimuli were made using luminescent patches (subtending approximately 0.2°) fixed to the experimenter at the neck, shoulders, elbows, wrists, hips, knees, and ankles. Live presentation was performed under blackout conditions within the laboratory. Video images were taken both of actors walking under strong diffuse lighting and under blackout conditions to give normal walking stimuli and the equivalent biological motion stimuli. The biological motion stimuli were then contrast thresholded to two luminance levels and finally contrast enhanced to black and white using a Fairlight Computer

Video instrument. Both these and the images under natural lighting conditions were stored on videodisc.

In addition to the small dot stimuli, stick figure representations were also used. These were generated in an analogous fashion to the biological motion stimuli but short strips of luminescent material were fixed between the articulation points. Gaps of similar size to the "dots" were left at the articulation points. These stick figures have more information than traditional biological motion stimuli since they give linkage structure but they do *not* have other form information (e.g., appearance of the face).

Control objects moving in the same directions as the walking/translating and biological motion stimuli were used. These were matched for size and like walking had nonrigid motion (e.g., curtains, lab-coats, hinged pieces of wood) and were moved at the same speed ($\pm 15\%$) and direction ($\pm 10^\circ$) as the walking stimuli. The responses of STPa cells to stimuli of the translating body will be reported elsewhere (Oram et al., in prep.). Three biological motion controls were also used. One was luminous dots moving nonrigidly under blackout conditions. Second, images of rigid translating dots were stored on videodisc. The third was a "jumbled" biological motion figure recorded on videodisc. The jumbled figure was made using a computer-based system (IRIS 3130, Silicon Graphics). The positions of the points of limb articulation were digitized for each of 24 video frames making one step cycle in 4 directions on a treadmill. To create normal motion sequences these were reanimated at 24 frames/sec and each point was allocated an additional translation vector to recreate walking motions with displacement (i.e., walking to the left). For jumbled figures the coordinates were moved in a random direction by a distance that was 30% of the initial head to floor height of the figure. The appropriate motion vector of each of the points was then added as was the translation vector. Therefore the resulting linkage structure was not changed but, when replayed, even though the overall translation and the individual component motions remained consistent with a walking stimulus, the image was no longer recognizable as a human figure. It is important to realize that the light points can still be "connected" by rigid limb elements and that the component motions were identical to the equivalent biological motion stimulus, but the relative lengths, the relative positions, and the relative motions of these elements were no longer humanoid.

The luminance values of the videodisc images were 0.2 cd/m^2 for the background, 3.0 cd/m^2 for the dots and stripes (see below), and 4.0 cd/m^2 for the natural image of a walking person. Under live conditions, the dot and stripe luminance was less than 0.1 cd/m^2 , whereas for natural images of a walking person the luminance was $1\text{--}4 \text{ cd/m}^2$.

The stimuli were viewed through either a liquid crystal shutter (Screen Print Technology) or a large aperture

camera shutter (Compur, 6.5 cm diameter). Both shutters had rise times of < 15 msec. The time at which the shutter became transparent or was fully open was also recorded. Subsequent analysis was linked to this, the true stimulus onset time. Each stimulus was presented five or more times in computer-controlled pseudo-random order. In addition, a "no stimulus" condition was also used, where only the LED and wall could be seen. This was used to assess background or spontaneous activity (SA) levels of the cells. Each trial consisted of a 0.5-sec warning tone, followed by a 1-sec stimulus presentation period. The intertrial interval was randomly varied between 0.5 and 5 sec. Motion of the stimulus was started before the presentation period and continued for a short duration afterward to ensure a smoothly moving presentation.

The isolated cells were tested using normal lighting and biological motion conditions. Tests were, at the least, of two body views and controls moving in one or two directions. If, under biological motion conditions, no response to the stick figures was found, then it was assumed that the dot figures would not elicit a response. Cells selectively responsive to body motions other than walking were tested for biological motion sensitivity with luminous patches only under laboratory blackout since appropriate images were not available on videodisc. Body motions used in this testing included rotation, crouching, and bowing. Cells were also tested for selectivity to the single limb movements present in the preferred stimulus. These tests involved the presentation of the arm or leg flexing and extending in isolation (i.e., rest of the body occluded from sight or visible but stationary). Cells found selective for these stimuli have been reported previously (e.g., leg, arm, or hand motion, see Perrett et al., 1985b, 1989a,b, 1990a,b; Mistlin & Perrett, 1990). The cells reported here, however, were unresponsive to individual limb movements.

Data Analysis

Analysis of the spike frequency data was performed on-line as a one-way ANOVA with each condition tested as a factor. The results of this analysis were used to guide subsequent testing. A cell was classified as selective for walking if there was a significant overall effect of conditions and one direction/body view combination was different ($p < 0.05$) from (1) control objects, (2) a second body view moving in the same direction, and (3) the same body view moving in a second direction. All the cells reported here were not found to be selective for single limb articulation but rather required whole body motion. Cells found to be selective for walking stimuli were then tested with both real walking and biological motion stimuli and subjected to off-line analysis. Off-line analysis for all cells took the form of two-way ANOVA, with the direction of motion as one factor and the stimulus type (natural, biological motion, con-

rol) as the second factor. A second two-way ANOVA was performed with body view as one factor and stimulus type (natural, biological motion) as the second. Significance for all statistical tests was taken at the 0.05 level. Post hoc testing of the ANOVAs was performed using the protected least significant difference (PLSD) test (Snedecor & Cochran, 1980).

Histological Reconstruction

After each recording session, frontal and lateral X-radiographs were taken. Reference lesions were made at the end of some of the tracks (10 μ A for 30 sec). In three monkeys additional reference markers were available from the injection sites of anatomical tracers (horseradish peroxidase and the fluorescent dyes, diamadino yellow and true blue) and in the fourth India ink was injected for further reference markers.

Following the last recording session, the subject was sedated with ketamine, then administered a lethal dose of barbiturate anesthetic. After transcardial perfusion with phosphate-buffered saline and 4% gluteraldehyde/paraformaldehyde fixative, the brain was removed and soaked in successively higher concentrations of sucrose solution or 2% dimethyl sulfoxide and 20% glycerol (Rosene, Roy, & Davis, 1986). Sections were taken every 500 μ m using standard techniques. The 3-D trajectory of each track was calculated from the X-radiograph coordinates. Cell positions along each track were then mapped onto the sections (see Harries & Perrett, 1991 for full details).

Acknowledgments

This research was funded by project grants from the M.R.C. (G8427112N), S.E.R.C. (GR/F 96723; GR/E 88257), United States O.N.R., and a Royal Society University Research Fellowship to D.P. We acknowledge the contribution of M. H. Harries, W. Dittrich, J. K. Hietanen, P. J. Benson, A. J. Chitty, A. J. Mistlin, and A. S. Head, who participated in some of the experiments, and to P. J. Benson for the programs allowing the jumbling of the biological motion stimuli. We thank the anonymous referees for their comments.

Reprint requests should be sent to Dr. D. I. Perrett, School of Psychology, University of St. Andrews, Scotland, KY16 9JU, U.K.

REFERENCES

- Albright, T. D. (1984). Direction and orientation selectivity of neurons in the visual area MT of the macaque. *Journal of Neurophysiology*, 52, 1106-1130.
- Albright, T. D., Desimone, R., & Gross, C. G. (1984). Columnar organization of directionally selective cells in visual area MT of the macaque. *Journal of Neurophysiology*, 51, 16-31.
- Andersen, R. A., & Siegel, R. M. (1989). Motion processing in the primate cortex. In G. M. Edelman, W. E. Gall, & W. M.

- Cowan (Eds.), *Signal and sense: Local and global order in perceptual maps* (pp. 163–184).
- Barlow, H. B., & Levick, W. R. (1965). The mechanism of directionally selective units in rabbit's retina. *Journal of Physiology (London)*, *178*, 477–504.
- Bassili, J. N. (1979). Emotion recognition: The role of facial movement and the relative importance of upper and lower areas of the face. *Journal of Personality and Social Psychology*, *37*, 2049–2058.
- Bennett, B. M., & Hoffman, D. D. (1985). The computation of structure from fixed-axis motion: Nonrigid structures. *Biological Cybernetics*, *51*, 293–300.
- 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.
- Boussaoud, D., Ungerleider, L. G., & Desimone, R. (1990). Pathways for motion analysis: Cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. *Journal of Comparative Neurology*, *296*, 462–495.
- Britten, K. H., Newsome, W. T., & Saunders, R. C. (1992). Effects of inferotemporal cortex lesions on form-from-motion discrimination in monkeys. *Experimental Brain Research*, *88*, 292–302.
- Brothers, L., & King, B. (1992). A neuroethological framework for the representation of minds. *Journal of Cognitive Neuroscience*, *4*, 107–118.
- Bruce, C. J., Desimone, R., & Gross, C. G. (1981). Visual properties on neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology*, *46*, 369–384.
- Campbell, R. (1992). The neuropsychology of lip-reading. *Philosophical Transactions of the Royal Society of London: Series B*, *335*, 39–45.
- Colby, C. L. (1991). The neuroanatomy and neurophysiology of attention. *Journal of Child Neurology*, *6*, S90-S118.
- Colby, C. L., & Miller, E. K. (1986). Eye movement related responses of neurons in superior temporal polysensory area of macaque. *Society of Neuroscience Abstracts*, *12*, 1.
- Cutting, J. E. (1978). Generation of synthetic male and female walkers through manipulation of a biomechanical invariant. *Perception*, *7*, 393–405.
- Cutting, J. E., & Kozlowski, L. T. (1977). Recognising friends by their walk: Gait perception without familiarity cues. *Bulletin of the Psychonomics Society*, *9*, 353–356.
- Cutting, J. E., Proffitt, D. R., & Kozlowski, L. T. (1978). A biomechanical invariant for gait perception. *Journal of Experimental Psychology*, *4*, 357–372.
- Cutting, J. E., Moore, C., & Morrison, R. (1988). Masking the motion of human gait. *Perception and Psychophysics*, *44*, 339–347.
- Dean, P. (1976). Effects of inferotemporal lesions on the behaviour of monkeys. *Psychological Bulletin*, *83*, 41–71.
- 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.
- De Yoe, F. A., & Van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. *Trends in the Neurosciences*, *11*, 219–226.
- Dittrich, W. D. (1993). Action categories and the perception of biological motion. *Perception*, *22*, 15–22.
- Duffy, C. J., & Wurtz, R. H. (1991a). Sensitivity of MST neurons to optic flow stimuli. I: A continuum of response selectivity to large-field stimuli. *Journal of Neurophysiology*, *65*, 1329–1345.
- Duffy, C. J., & Wurtz, R. H. (1991b). Sensitivity of MST neurons to optic flow stimuli. II: Mechanisms of response selectivity revealed by small-field stimuli. *Journal of Neurophysiology*, *65*, 1346–1359.
- Ekman, P., & Friesen, W. V. (1976). *Pictures of facial affect*. Palo Alto, CA: Consulting Psychologists Press.
- Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*, 1–47.
- Fox, R., & McDaniel, C. (1982). The perception of biological motion by human infants. *Science*, *218*, 486–487.
- Goddard, N. H. (1992). *The perception of articulated motion: Recognizing moving light displays*. Ph.D. thesis, University of Rochester, New York.
- Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in the Neurosciences*, *15*, 20–25.
- Gross, C. G. (1992). Representation of visual stimuli in inferior temporal cortex. *Philosophical Transactions of the Royal Society of London: Series B*, *335*, 3–10.
- Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the monkey. *Journal of Neurophysiology*, *35*, 96–111.
- Harries, M. H., & Perrett, D. I. (1991). Modular organization of face processing in temporal cortex: Physiological evidence and possible anatomical correlates. *Journal of Cognitive Neuroscience*, *3*, 9–24.
- Hasselmo, M. E., Rolls, E. T., Baylis, G. C., & Nalwa, V. (1989). Object-centered encoding by face-selective neurons in the cortex in the superior temporal sulcus of the monkey. *Experimental Brain Research*, *75*, 417–429.
- Hietanen, J. K., & Perrett, D. I. (1993). Motion sensitive cells in the macaque superior temporal polysensory area: I. Lack of response to the sight of the monkey's own hand. *Experimental Brain Research*, *93*, 117–128.
- Hikosaka, K., Iwai, E., Saito, H.-A., & Tanaka, K. (1988). Polysensory properties of neurons in the anterior bank of the caudal superior temporal sulcus of the macaque monkey. *Journal of Neurophysiology*, *60*, 1615–1637.
- Hildreth, E. C., & Koch, C. (1987). The analysis of visual motion: From computational theory to neuronal mechanisms. *Annual Review of Neuroscience*, *10*, 477–533.
- Hoffman, D. D., & Flinchbaugh, B. E. (1982). The interpretation of biological motion. *Biological Cybernetics*, *42*, 195–204.
- Humphreys, G. W., Donnelly, N., & Riddoch, M. J. (1993). Expression is computed separately from facial identity and it is computed separately for moving and static faces: Neuropsychological evidence. *Neuropsychologia*, *31*, 173–181.
- Jansson, G., & Johansson, G. (1973). Visual perception of bending motion. *Perception*, *2*, 321–326.
- Johansson, G. (1973). Visual perception of biological motion and a model for its analysis. *Perception and Psychophysics*, *14*, 201–211.
- Johansson, G. (1976). Spatio-temporal differentiation and integration in visual motion perception. *Psychological Research*, *38*, 379–393.
- Koenderink, J. J., & van Doorn, A. J. (1979). The internal representation of solid shape with respect to vision. *Biological Cybernetics*, *32*, 211–216.
- Komatsu, H., & Wurtz, R. H. (1988a). Relation of cortical areas MT and MST to pursuit eye movements. I. Localization and visual response properties of neurons. *Journal of Neurophysiology*, *60*, 580–603.
- Komatsu, H., & Wurtz, R. H. (1988b). Relation of cortical areas MT and MST to pursuit eye movements. III. Interaction with full-field visual stimulation. *Journal of Neurophysiology*, *60*, 621–644.
- Kozlowski, L. T., & Cutting, J. E. (1977). Recognizing the sex

- of a walker from a dynamic point-light display. *Perception and Psychophysics*, 21, 575–580.
- Lappin, J. S., Doner, J. F., & Kottas, B. (1980). Minimal conditions for the visual detection of structure and motion in three dimensions. *Science*, 209, 717–719.
- Lee, H. J., & Chen, Z. (1985). Determination of 3D human body postures from a single view. *Computer Vision, Graphics and Image Processing*, 30, 148–168.
- Leung, M. K., & Yang, Y. H. (1987). A region-based approach for human body motion analysis. *Pattern Recognition*, 20, 321–339.
- Lowe, D. G. (1987). *Perceptual organization and visual recognition*. Boston: Kluwer Academic Publishers.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco, CA: Freeman.
- Marr, D., & Nishihara, H. K. (1978). Representation and recognition of the spatial organization of three dimensional shapes. *Proceedings of the Royal Society of London: Series B*, 200, 269–294.
- Marr, D., & Vaina, L. (1982). Representation and recognition of the movements of shapes. *Proceedings of the Royal Society of London: Series B*, 214, 501–524.
- 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.
- Mikami, A., Newsome, W. T., & Wurtz, R. H. (1986a). Motion sensitivity in macaque visual cortex. I: Mechanisms of direction and speed selectivity in extrastriate area MT. *Journal of Neurophysiology*, 55, 1308–1327.
- Mikami, A., Newsome, W. T., & Wurtz, R. H. (1986b). Motion selectivity in macaque visual cortex. II: Spatiotemporal range of directional interactions in MT and V1. *Journal of Neurophysiology*, 55, 1328–1339.
- Mishkin, M., Ungerleider, L. G., & Macko, K. A. (1983). Object vision and spatial vision: Two cortical pathways. *Trends in the Neurosciences*, 6, 414–417.
- Mistlin, A. J., & Perrett, D. I. (1990). Visual and somatosensory processing in the macaque temporal cortex: The role “expectation.” *Experimental Brain Research*, 82, 437–450.
- Oram, M. W., & Perrett, D. I. (1993). Neural processing of biological motion in the macaque temporal cortex. S.P.I.E. conference proceedings, Asilomar, June 1993, in press.
- Oram, M. W., Perrett, D. I., & Hietanen, J. K. (1993). Directional tuning of motion sensitive cells in the anterior superior temporal polysensory area (STPa) of the macaque. *Experimental Brain Research*, 97, 274–294.
- Perrett, D. I., Rolls, E. T., & Caan, W. (1982). Visual neurons responsive to faces in the monkey temporal cortex. *Experimental Brain Research*, 47, 329–342.
- Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1984). Neurons responsive to faces in the temporal cortex: Studies of functional organization sensitivity to identity and relation to perception. *Human Neurobiology*, 3, 197–208.
- Perrett, D. I., Smith, P. A. J., Potter, D. D., Mistlin, A. J., Head, A. S., Milner, A. D., & Jeeves, M. A. (1985a). Visual cells in the temporal cortex sensitive to face view and gaze direction. *Proceedings of the Royal Society of London: Series B*, 223, 293–317.
- Perrett, D. I., Smith, P. A. J., Mistlin, A. J., Chitty, A. J., Head, A. S., Potter, D. D., Broennimann, R., Milner, A. D., & Jeeves, M. A. (1985b). Visual analysis of body movements by neurons in the temporal cortex of the macaque monkey: A preliminary report. *Behavioural Brain Research*, 16, 153–170.
- Perrett, D. I., Mistlin, A. J., Chitty, A. J., Smith, P. A. J., Potter, D. D., Broennimann, R., & Harries, M. (1988). Specialized face processing and hemispheric asymmetry in man and monkey: Evidence from single unit and reaction time studies. *Behavioural Brain Research*, 29, 245–258.
- Perrett, D. I., Mistlin, A. J., Harries, M. H., & Chitty, A. J. (1989a). Understanding the visual appearance and consequence of hand actions. In *Vision and action: The control of grasping* (pp. 163–180). Norwood, NJ: Ablex Publishing.
- Perrett, D. I., Harries, M. H., Bevan, R., Thomas, S., Benson, P. J., Mistlin, A. J., Chitty, A. J., Hietanen, J. K., & Ortega, J. E. (1989b). Frameworks of analysis for the neural representation of animate objects and actions. *Journal of Experimental Biology*, 146, 87–114.
- Perrett, D. I., Harries, M. H., Benson, P. J., Chitty, A. J., & Mistlin, A. J. (1990a). Retrieval of structure from rigid and biological motion: An analysis of the visual response of neurons in the macaque temporal cortex. In A. Blake, & T. Troscianko (Eds.), *AI and the Eye* (pp. 181–201). Chichester, England: John Wiley.
- Perrett, D. I., Harries, M. H., Chitty, A. J., & Mistlin, A. J. (1990b). Three stages in the classification of body movements by visual neurons. In H. B. Barlow, C. Blakemore, & M. Weston-Smith (Eds.), *Images and understanding* (pp. 94–107). Cambridge, England: Cambridge University Press.
- Perrett, D. I., Oram, M. W., Harries, M. H., Bevan, R., Hietanen, J. K., Benson, P. J., & Thomas, S. (1991). Viewer-centred and object-centred coding of heads in the macaque temporal cortex. *Experimental Brain Research*, 86, 159–173.
- Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London: Series B*, 335, 23–30.
- Poizner, H., Bellugi, U., & Lutes-Driscoll, V. (1981). Perception of American Sign Language in dynamic point-light displays. *Journal of Experimental Psychology and Human Perception and Performance*, 7, 430–440.
- Proffitt, D. R., Bertenthal, B. I., & Roberts, Jr., R. J. (1984). The role of occlusion in reducing multistability in moving point-light displays. *Perception and Psychophysics*, 36, 315–323.
- Rashid, R. F. (1980). Towards a system for the interpretation of moving light displays. *IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-2*, 574–581.
- Rodman, H. R., & Albright, T. D. (1989). Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). *Experimental Brain Research*, 75, 53–64.
- Rosene, D. L., Roy, N. J., & Davis, B. J. (1986). A cryoprotection method that facilitates cutting frozen sections of whole monkey brains for histological and histochemical processing without freezing artifact. *Journal of Histochemistry Cytochemistry*, 34, 1301–1315.
- Saito, H., Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y., & Iwai, E. (1986). Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *Journal of Neuroscience*, 6, 145–157.
- Saito, H., Tanaka, K., Isono, H., Yasuda, M., & Mikami, A. (1989). Directionally selective response of cells in the middle temporal area (MT) of the macaque monkey to the movement of equiluminous opponent color stimuli. *Experimental Brain Research*, 75, 1–14.
- Seibert, M., & Waxman, A. M. (1991). Learning aspect graph representations from view sequences. In D. S. Touretzky (Ed.), *Advances in neural network information processing systems* (Vol. 2, pp. 258–265). San Mateo, CA: Morgan Kaufmann.
- Seibert, M., & Waxman, A. M. (1992a). Adaptive 3D object recognition from multiple views. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14, 107–124.
- Seibert, M., & Waxman, A. M. (1992b). Learning and recogniz-

- ing 3D objects from multiple views in a neural system. In H. Wechsler (Ed.), *Neural networks for perception* (Vol. 1, part 2, pp. 426–444). San Diego: Academic Press.
- Seltzer, B., & Pandya, D. N. (1978). Afferent cortical connections and architectonics of the superior temporal sulcus and surrounding cortex in the rhesus monkey. *Brain Research*, *149*, 1–24.
- Siegel, R. M., & Andersen, R. A. (1986). Motion perceptual deficits following ibotenic acid lesions of the middle temporal area in the behaving rhesus monkey. *Society of Neuroscience Abstracts*, *12*, 1183.
- Siegel, R. M., & Andersen, R. A. (1990). The perception of structure from visual motion in monkey and man. *Journal of Cognitive Neuroscience*, *2*, 306–319.
- Snedecor, G. W., & Cochran, G. W. (1980). *Statistical methods*. Ames: Iowa State University Press.
- Snowden, R. J., Treue, S., Erickson, R. G., & Andersen, R. A. (1991). The response of area MT and V1 neurons to transparent motion. *Journal of Neuroscience*, *11*, 2768–2785.
- Snowden, R. J., Treue, S., & Andersen, R. A. (1992). The response of neurons in areas V1 and MT of the alert rhesus monkey to moving random dot patterns. *Experimental Brain Research*, *88*, 389–400.
- Sugie, N., & Kato, K. (1987). A computational model of biological motion perception. *IEEE Montech, Conference on Biomedical Technologies, November 1987*, 140–143.
- Sugihara, K., & Sugie, N. (1984). Recovery of rigid structure from orthographically projected optical-flow. *Computer Vision, Graphics and Image Processing*, *27*, 309–320.
- Sumi, S. (1984). Upside-down presentation of the Johansson moving light-spot pattern. *Perception*, *13*, 283–286.
- Tanaka, K., & Saito, H-A. (1989). Analysis of motion of the visual field by direction, expansion/contraction and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. *Journal of Neurophysiology*, *62*, 626–641.
- Tanaka, K., Hikosaka, K., Saito, H-A., Yukie, M., Fukada, Y., & Iwai, E. (1986). Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. *Journal of Neuroscience*, *6*, 134–144.
- Tanaka, K., Fukada, Y., & Saito, H. (1989). Underlying mechanisms of the response specificity of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey. *Journal of Neurophysiology*, *62*, 642–656.
- Tanaka, K., Saito, H., Fukada, Y., & Moriya, M. (1991). Coding visual images of objects in the inferotemporal cortex of the macaque monkey. *Journal of Neurophysiology*, *66*, 170–189.
- Torre, V., & Poggio, T. (1978). A synaptic mechanism possibly underlying directional selectivity to motion. *Proceedings Royal Society London: Series B*, *202*, 409–416.
- Tovee, M. J., & Rolls, E. T. (1993). The temporal responses of translation-invariant neurons in the primate temporal visual cortex. *Perception*, *22 Suppl.*, 141.
- Ullman, S. (1979). The interpretation of structure from motion. *Proceedings of the Royal Society of London: Series B*, *203*, 405–426.
- Ungerleider, L. G., & Mishkin, M. (1982). Two cortical visual systems. In D. J. Ingle, M. A. Goodale, & R. J. W. Mansfield (Eds.), *Analysis of visual behaviour* (pp. 549–586). Cambridge, MA: MIT Press.
- Vaina, L. M., Lemay, M., Bienfang, D. C., Choi, A. Y., & 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.
- Webb, J., & Aggarwal, J. (1982). Structure from motion of rigid and jointed objects. *Artificial Intelligence*, *19*, 107–130.
- Young, M. P. (1992). Objective analysis of the topological organization of the primate cortical visual system. *Nature (London)*, *358*, 152–155.
- Young, M. P., & Yamane, S. (1992). Sparse population coding of faces in the inferotemporal cortex. *Science*, *256*, 1327–1331.
- Zeki, S. M. (1974). Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *Journal of Physiology (London)*, *236*, 549–573.
- Zihl, J. D., Von Cramon, D., & Mai, N. (1983). Selective disturbance of vision after bilateral brain damage. *Brain*, *106*, 313–340.