



# Cells in monkey STS responsive to articulated body motions and consequent static posture: a case of implied motion?

Tjeerd Jellema<sup>a,\*</sup>, David I. Perrett<sup>b</sup>

<sup>a</sup> Psychological Laboratory, Helmholtz Research Institute, Utrecht University, Utrecht 3584 CS, The Netherlands

<sup>b</sup> School of Psychology, University of St. Andrews, Fife KY16 9JU, UK

Received 25 April 2003; accepted 20 June 2003

## Abstract

We show that populations of visually responsive cells in the anterior part of the superior temporal sulcus (STSa) of the macaque monkey code for the sight of both specific articulated body actions and the consequent articulated static body postures. We define articulated actions as actions where one body part (e.g. a limb or head) moves with respect to the remainder of the body which remains static; conversely non-articulated actions are actions where the equivalent body parts do not move with respect to each other but move as one. Similarly, articulated static body postures contain a torsion or rotation between parts, while non-articulated postures do not. Cells were tested with the sight of articulated and non-articulated actions followed by the resultant articulated or non-articulated static body postures. In addition, the static body postures that formed the start and end of the actions were tested in isolation. The cells studied did not respond to the sight of non-articulated static posture, which formed the starting-point of the action, but responded vigorously to the articulated static posture that formed the end-point of the action. Other static postures resembling the articulated end-point posture, but which were in a more relaxed muscular state (i.e. non-articulated), did not evoke responses. The cells did not respond to body actions that were less often associated with the effective static articulated postures. Our results suggest that the cells' responses were related to the implied action rather than the static posture per se. We propose that the neural representations in STSa for *actual* biological motion may also extend to biological motion *implied* from static postures. These representations could play a role in producing the activity in the medial temporal/medial superior temporal (V5(MT)/MST) areas reported in fMRI studies when subjects view still photographs of people in action.

© 2003 Elsevier Ltd. All rights reserved.

**Keywords:** Body posture; Visual coding; Biological motion; Implied motion

## 1. Introduction

Actions performed by (non-)human primates typically involve articulation, in which one or more body part(s), e.g. the head or a limb, move with respect to the remainder of the body. Purely non-articulated actions, in which the body moves solidly, are rare in the animal kingdom (an example is a hawk diving down onto a prey). To understand an articulated action performed by another individual, we do not necessarily have to witness the entire action sequence. A single momentary view is often enough to identify the action and possibly also the goal of the action. This is a very useful capacity since it allows us to understand other's actions in situations where the other is intermittently occluded from view, or when we only get one glimpse.

A brain structure particularly relevant for processing visual object motion is the medial temporal/medial superior

temporal (V5(MT)/MST) complex. The V5(MT)/MST complex plays a primary role in the analysis of the direction and speed of moving objects in the visual world, as shown by a plethora of single-cell studies in monkeys (Britten, Shadlen, Newsome, & Movshon, 1992; Britten, Newsome, Shadlen, Celebrini, & Movshon, 1996; Shadlen, Britten, Newsome, & Movshon, 1996; Dubner & Zeki, 1971; Maunsell & van Essen, 1983) and fMRI studies in humans (Smith, Greenlee, Singh, Kraemer, & Henning, 1998; Watson et al., 1993). In monkeys, the processing of *biological* motion (i.e. the natural articulation of bodies during actions) involves additional areas anterior to the V5(MT)/MST complex, including a region situated in the anterior part of the superior temporal sulcus (STSa), corresponding to area STPa (Bruce, Desimone, & Gross, 1981). Single-cell studies in the macaque monkey revealed that STSa cells selectively respond to visual cues derived from (parts of) the whole static or moving body (Perrett, Hietanen, Oram, & Benson, 1992; Jellema & Perrett, 2002; Perrett et al., 1985a,b; Perrett et al., 1989). Many STSa cells integrate information about the form and

\* Corresponding author. Tel.: +31-30-253-4368; fax: +31-30-253-4511.  
E-mail address: t.jellema@fss.uu.nl (T. Jellema).

motion of animate objects (Oram & Perrett, 1996). Subpopulations of STSa cells are sensitive to the spatial position of the object (Baker, Keysers, Jellema, Wickers, & Perrett, 2000). These findings lend support to the idea that this part of temporal cortex acts as an interface between the dorsal and the ventral streams of visual processing (Karnath, 2001). Collectively, the response properties suggest that STSa contributes to the visual recognition and understanding of other's actions (Jellema, Baker, Wicker, & Perrett, 2000; Jellema, Baker, Oram, & Perrett, 2002; Emery & Perrett, 1994; Jellema & Perrett, 2002). In humans, imaging studies showed activation of the STS in response to aversion of eye gaze (Hoffman & Haxby, 2000), and to movement of the hands (Grafton, Arbib, Fadiga, & Rizzolatti, 1996; Rizzolatti et al., 1996), the face (Puce, Allison, Bentin, Gore, & McCarthy, 1998), and to point light displays of whole body motion (Bonda, Petrides, Ostry, & Evans, 1996) and meaningful actions (Decety et al., 1997); see (Allison, Puce, & McCarthy, 2000 for a review). In line with these findings, a direct comparison between articulated and non-articulated human motion showed a greater response to articulated human motion in the STS (Beauchamp, Lee, Haxby, & Martin, 2003). In addition, the right lateral occipitotemporal cortex is activated by static images of many different parts of the human body, but not by whole faces (Downing, Yiang, Shuman, & Kanwisher, 2001).

Recently, two separate fMRI studies suggested that the V5(MT)/MST complex is also involved in the processing of *implied* motion of objects (Kourtzi & Kanwisher, 2000; David & Senior, 2000; Senior et al., 2000). In the implied motion condition, a still photograph of an object in motion is shown (e.g. a snapshot of an athlete running, or of a cup falling from a shelf). This was a startling finding, for, up to then, it was believed that V5(MT)/MST could only be activated by perceiving actual motion, or by the illusion that motion occurs, such as in apparent motion (Goebel, Khorram-Sefat, Muckli, Hacker, & Singer, 1998), illusory motion (Zeki, Watson, & Frackowiak, 1993), imagined motion (Cohen et al., 1995) and the motion after-effect (Tootell et al., 1995), all of which excite V5(MT)/MST. The recent evidence suggests that V5(MT)/MST can be activated in the absence of any direct visual motion experience.

The activity in V5(MT)/MST to implied (non-)biological object motion in static images suggests a top-down influence onto V5(MT)/MST, since the object and context may first need to be identified, before the likely type of movements associated with the object in that context can be determined. These object recognition and knowledge related processes suggest an involvement of the ventral recognition stream (Gross, Rocha-Miranda, & Bender, 1972; Milner & Goodale, 1995; Ungerleider & Mishkin, 1982). Sensitivity to associations between image form and motion could form the basis of the ability of the nervous system to retrieve likely motion given entirely static images. In this context, we have studied the sensitivity of STSa cells to the associations between particular body postures

and body movements. Populations of cells in this region are already known to be involved in the processing of the view of separate body components (Wachsmuth, Oram, & Perrett, 1994) and articulation of body components (Perrett et al., 1985b). There is also evidence that form and motion signals are extensively integrated within the STSa (Oram & Perrett, 1996). This led us to speculate that the conjoint cell sensitivity to particular posture and articulation present in actual biological motion might underlie representation of implied biological motion. We therefore compared STS cell responses to static postures and to related actions.

In our experiments cells were tested with: (a) a variety of articulated and non-articulated body actions, (b) the starting- and end-postures of these actions presented 'in isolation', and (c) other static postures. Non-articulated actions were chosen such that the body part in question moved in a similar way to that of the articulated action. For example: head rotating toward observer with the trunk remaining stationary (articulated action) versus head and trunk rotating simultaneously toward the observer (non-articulated action). Our main hypothesis was that STSa cells sensitive to a particular articulated action would also be sensitive to an associated static posture when presented in isolation (i.e. in absence of actual movement). The nature of any association between sensitivity to posture and sensitivity to action could provide insight into the neural basis of phenomena such as implied motion.

## 2. Method

### 2.1. Subjects and physiological procedures

The experiments were performed on two awake rhesus macaque monkeys (*Macaca mulatta*, age 4–6 years). A detailed description of the surgical procedures can be found elsewhere (e.g. Oram & Perrett, 1996). Animal care and experimental procedures were performed in accordance with UK Home Office guidelines.

### 2.2. Recording

Spikes were captured online onto a PC (CED1401plus, Cambridge Electronic Design, UK). Additionally, spikes were stored on an audio track of a HiFi videotape recorder. The stimulus events (seen from the subject's perspective) were recorded with a video camera, and stored simultaneously on the video track of the same tape. Eye movements were recorded with a second (infra-red sensitive) camera mounted onto the primate chair. The signals from the two cameras were integrated (Panasonic VHS video mixer, WJAVE7) prior to recording. The signal from the eye camera was also recorded separately on a second video tape recorder, synchronised with a time-code generator and frame counter (VITC Horita VG50), for off-line analysis of eye position (Iviev, Sensomotoric Instruments, Germany). Trials at which the eyes were closed were discarded. Analysis

of the eye positions indicated no relationship between the pattern of fixation and trial type or response magnitude.

### 2.3. Stimuli

Visual stimuli consisted of a wide variety of (non-) articulated body actions. When a cell was found to respond specifically to an articulated action, three further conditions were employed: (1) Similar, but non- or less articulated, actions were presented. For example if the articulated action consisted of a rotation of the head from right profile head view to face view, with the trunk remaining static in the right profile view, then the non-articulated action would be a similar head rotation from right profile to face view but with the trunk rotating along with the head (no torsion between head and trunk). (2) The non-articulated starting-point (e.g. right profile head, right profile chest) and the articulated end-point of the action (e.g. front face view head, right profile chest) were presented in isolation (i.e. presented immediately following opening of the shutter). (3) Static, non-articulated postures similar to the effective articulated posture, were presented (e.g. front view head, front view chest). Some movements do not have a natural end-posture, in which case the end-posture was defined as the point in the trajectory of an action at which the cell stopped responding if the action continued. For example a cell might respond during rotation of the head toward the observer but would cease to respond to rotation once the head began to turn away from the observer. Stimuli were presented either on film projected onto a screen at life size, or live (only for human stimuli) from behind a fast rise-time liquid crystal shutter (aperture 20 cm × 20 cm at a distance of 15 cm). Video stimuli consisted of both monkeys and people; most cells do not discriminate between these two species (Perrett et al., 1985b). In some cases a mechanical shutter with a larger aperture was used to provide the subject with a wide field of view. The live stimuli were shown at 1.5–4 m distance from the subject. Five to twelve repetitions were tested per stimulus condition in pseudo random order. Control stimuli included objects of comparable size moved in comparable ways (e.g. a box attached onto the end of a stick, making rotational motions, served as a control for head rotations). The best controls, however, were formed by the non-articulated actions, which resembled the articulated actions more closely than any non-living control object could ever do.

### 2.4. Data analysis

Spikes from multiple single neurons were routinely discriminated on- or off-line using template matching (Spike2, Cambridge Electronic Design, UK). Spike counts were performed during presentation of the stimuli, usually at 1 s intervals. Presentation of the body actions lasted typically 700–800 ms (indicated by dotted lines in Figs. 1–3). Cell responses were analysed using ANOVAs, Newman–Keuls post-hoc testing and *t* tests (significance level at  $P < 0.05$ ).

In the ANOVAs, the factor time was used as a repeated measures variable, usually with 2–4 levels consisting of consecutive seconds of the recording period following opening of the shutter.

### 2.5. Cell localisation

A detailed description can be found elsewhere (Jellema et al., 2000). At completion of each experiment, frontal and lateral X-ray photographs were taken with the electrode still in place, to locate the electrode and the recorded cells with respect to specific bone landmarks. The subject was then sedated and given a lethal dose of anaesthetic. After transcardial perfusion the brain was removed, coronal sections (25  $\mu$ m) were cut, photographed and stained. The X-ray photographs were aligned with the histological sections to determine the cell locations (accuracy  $\approx 1$  mm). Histological reconstruction from one monkey is shown in Fig. 4. All cells localized were found in the upper bank of the STSa, between 10 and 19 mm anterior to the inter-aural-line.

## 3. Results

### 3.1. Single-cell responses to body postures and actions

In a sample of 272 visually responsive STSa cells, 208 cells (76%) showed maximal responses when the visual stimulus was a conspecific or a human, performing a particular action or assuming a particular posture. The other 64 cells (64/272, 24%) responded to various visual stimuli and motions and were not further studied. The majority of the cells (125/208, 60%) responded exclusively to body actions, and failed to respond to static body postures. The other 83 cells (83/208, 40%) did show responses to presentation of specific static body postures.

Forty-four cells of the latter category were subjected to detailed testing by presenting a range of body actions and postures. We found 24 cells (24/44, 55%) that specifically responded to articulated actions. The collection of articulated body actions we report here consisted predominantly of rotations of the head with respect to the trunk, or of rotations of the head and upper body with respect to the lower body, in either the horizontal plane (toward or away from the observer,  $n = 14$ ) or the vertical plane (upward or downward,  $n = 8$ ). The head rotations were performed by the experimenter while sitting on a mobile chair (on wheels), presented live or on video. Rotation of this mobile chair with the experimenter sitting on it produced the non-articulated, or whole body, rotations in the horizontal plane. For the head rotations in the vertical plane, the corresponding ‘whole body’ actions consisted of bending actions of the upper body and head conjointly. The other 20 cells (20/44, 43%) did not discriminate between whole body motions and articulations of body parts. These cells responded, e.g. to an anti-clockwise rotation of the head over the left shoulder toward the observer, but also

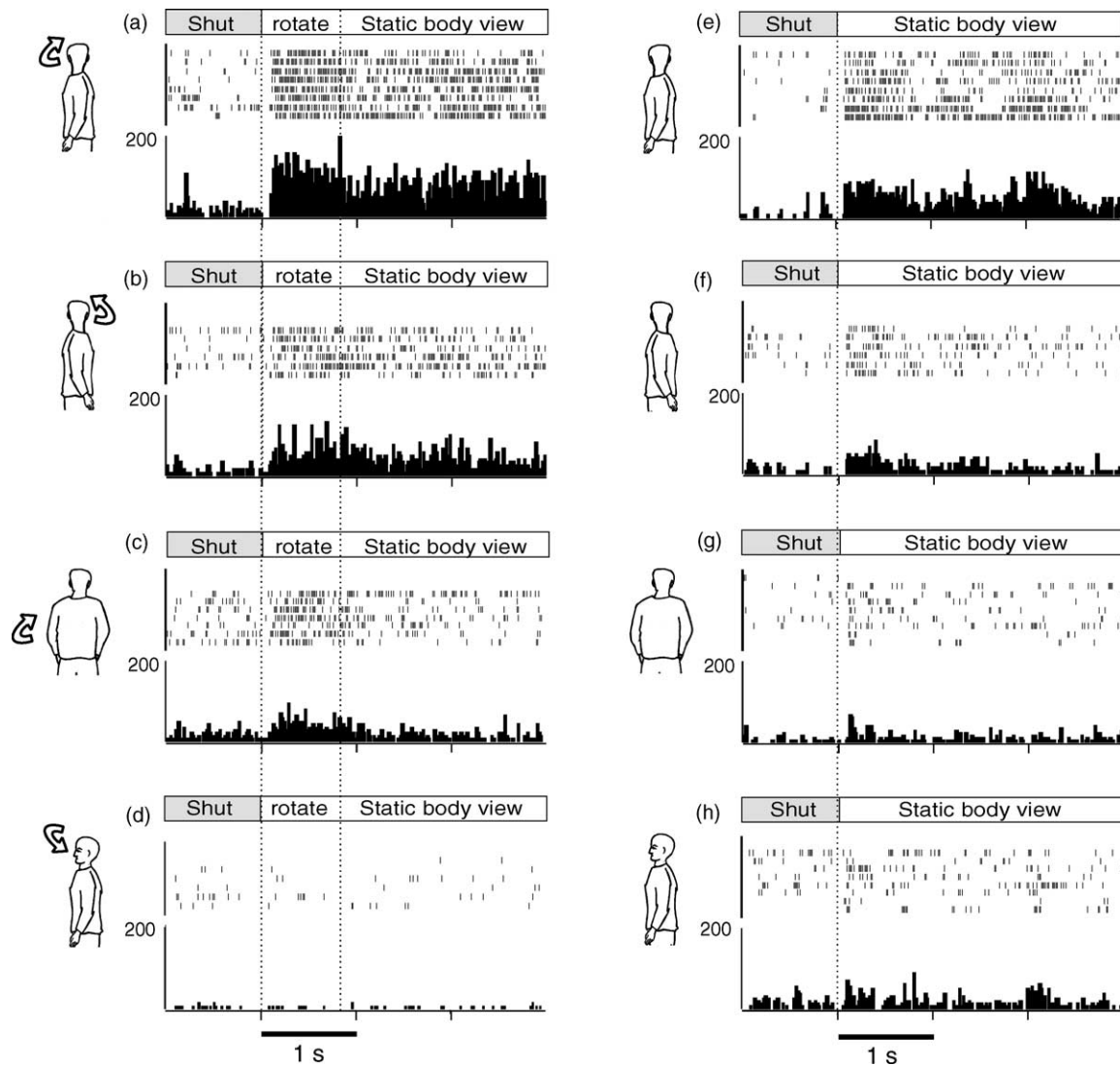


Fig. 1. Cell response to articulation of the head. During the first second of each record the shutter was closed ('shut') and the rastergrams show the spontaneous cell activity. During 3 s following opening of the shutter, the visual stimulus (direction of articulation denoted by arrow and final body posture denoted by figure) is presented. Left-hand column (a–d): body actions followed by static postures; right-hand column (e–g): static body views. (a) The cell responded maximally to the sight of the head rotating over the right shoulder away from the observer; (b) the mirror-image of the action presented in (a) evoked a significantly smaller response; (c) whole body rotation from left profile view to back view evoked a small response; (d) the same starting posture as (a) but rotating the head in the opposite direction (i.e. over the left shoulder) failed to evoke a response; (e) the static view of the left profile of the body with the head rotated away from the observer over the right shoulder, which is the end-posture of the action in (a), evoked a considerable response. Other static body views such the mirror-image (f), the non-articulated static back view of body and head (g), and the non-articulated static left profile view of body and head (h), evoked significantly smaller responses.

to an anti-clockwise rotation of the head and trunk toward the observer (while sitting on a mobile chair). Some cells in this category responded exclusively to the whole body motion.

Responses to the articulated actions were often highly specific. For example the cell in Fig. 1 responded vigorously to an articulated action consisting of a horizontal head rotation to the right shoulder, starting at the left profile view and ending with the back view of the head. The trunk did not rotate along with the head but remained static throughout (in the left profile view; Fig. 1a). Other body actions resembling the

one in Fig. 1a, such as rotation of the head from the right profile view to the back view (i.e. the mirror-image of the action in (a); Fig. 1b), and rotation of the whole of the upper body (head and trunk) from left profile view to back view (Fig. 1c, rotation of mobile chair), produced significantly smaller responses. Rotation of the head from back view to the left profile, i.e. the reverse of the action in Fig. 1a, did not evoke a response (Fig. 1d).

When the presentation of the static articulated end-posture was continued following the action, the cells continued to respond, often for several seconds, as illustrated in Fig. 1a



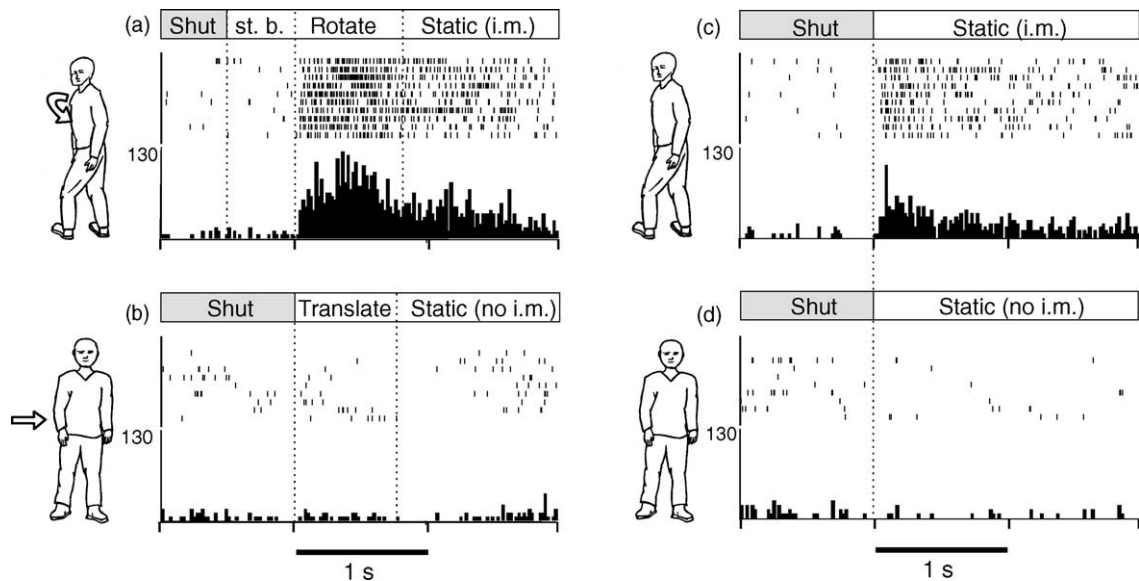


Fig. 2. Cell response to articulation of the upper body in horizontal plane. Visual stimuli are presented for 3 s following opening of the shutter. Left-hand column (a and b): body actions followed by static postures; right-hand column (c and d): static body views. (a) A rotation of the upper body with respect to the static lower body toward the observer evoked maximal responses; (b) horizontal motion of the body failed to evoke a response; (c) the articulated end-posture of the action in (a) presented in isolation evoked a considerable response; (d) the static, non-articulated, front view of body and head failed to evoke a response.

(right half of panel). A surprising finding was that the presentation of the articulated static end-posture in isolation (following opening of the shutter) also evoked large responses (Fig. 1e). The mirror-image of the posture in Fig. 1e (left profile trunk with back view of head; Fig. 1f) produced a significantly smaller response. Other static postures that represented a body 'at rest', such as the non-articulated back view of the trunk and head (Fig. 1g), failed to produce a response. The non-articulated starting posture of the action (left profile view of body and head; Fig. 1h) also did not produce a response.

Two-way ANOVA of Condition (eight levels: a–h)  $\times$  Time (four levels: the 1st, 2nd, 3rd and the 4th seconds) showed a main effect of Condition ( $F(7, 46) = 22.3, P < 0.00001$ ), a main effect of Time ( $F(3, 138) = 46.4, P < 0.00001$ ), and an interaction of Condition and Time ( $F(21, 138) = 8.7, P < 0.000001$ ). Post-hoc testing for the first second following shutter opening, showed that the response during the effective action (head rotating away from the observer, Fig. 1a) was significantly larger than the responses in all other conditions in the equivalent time-period (Newman–Keuls,  $P < 0.00005$ ). The response to the static articulated body posture following shutter opening (Fig. 1e) was significantly larger than the responses to the other static postures (Fig. 1f–h) ( $P < 0.05$ ).

All cells of this population (24/44) discriminated sharply between static views of the body with and without articulation, which is quite remarkable given the resemblance of the two types of postures. Here, a non-articulation posture refers to the head, chest and pelvis aligned and oriented in the same

direction typical of an at rest posture. It should be noted that the cell responses cannot be explained by sensitivity to the view of just one body segment. For example the response in Fig. 1e could in principle be due to (1) the back view of the head, or (2) the left profile view of the trunk. However, the virtual absence of response to the non-articulated back view of the body (Fig. 1g) shows that the back view of the head is not sufficient in itself to evoke a response. Likewise, the left profile view of the trunk is not sufficient as shown by the lack of response in Fig. 1h. Apparently the cell required torsion between trunk and head to respond. Indeed, the response selectivity was even stricter as only a head rotated over the right shoulder was effective, while a rotation over the left shoulder was much less effective (Fig. 1f). The cell in Fig. 1 used a viewer-centred (not object-centred) frame of reference (Perrett et al., 1989), since the same 90° clockwise head rotation seen from a different perspective (e.g. the front view, torso: front view; head: left profile view) failed to produce a response (data not shown).

Another example of a cell responding specifically to a particular articulated action is given in Fig. 2. The effective action for this cell consisted of a rotation of the upper body with respect to the static lower body (hips and legs) toward the observer in the horizontal plane (Fig. 2a). Rotation of just the head was ineffective (data not shown). This cell also used a viewer-centred frame of reference; the identical change in posture seen from a different perspective did not excite the cell. The presentation of other bodily actions, e.g. the translation of the upper body from left to right (Fig. 2b), did not excite the cell. The opposite action of that shown

in Fig. 2a, i.e. rotation of the upper body away from the observer, also failed to evoke a response (data not shown). Presentation of just the static articulated end-posture produced a considerable response (Fig. 2c), while the static

non-articulated front view of the body evoked no response (Fig. 2d). Again, the response to the static articulated posture (Fig. 2a) cannot be explained by response sensitivity to either of the two main body components present in the image: the face view did not evoke a response in combination with the frontal body view (d), nor did the non-articulated back and profile views of the body evoke a response (data not shown). Two-way ANOVA of Condition (four levels: a–d)  $\times$  Time (three levels: the 1st, 2nd and 3rd second, following opening of the shutter) showed a main effect of Condition ( $F(3, 33) = 91.6, P < 0.00001$ ), a main effect of Time ( $F(2, 66) = 80.7, P < 0.00001$ ), and an interaction for Condition and Time ( $F(6, 66) = 42.0, P < 0.00001$ ). Post-hoc testing for the first second following shutter opening showed that the response in (a) was larger than that in all other conditions ( $P < 0.0002$ ) and that the cell response discriminated between static postures (c) and (d), ( $P < 0.0002$ ).

### 3.2. Population responses

The population responses in the four main stimulus conditions are shown in Fig. 3. These conditions are: (1) the articulated body action the cell was tuned to, followed by the articulated static end-posture; (2) similar, but non-articulated, ‘whole’ body actions, followed by the non-articulated static end-posture; (3) the static articulated end-posture of the effective action presented in isolation; (4) the static non-articulated end-posture presented in isolation. Population responses in conditions 1 and 2 are depicted in Fig. 3a, and reveal a much greater effectiveness of the articulated action compared to the non-articulated action, both during the action phase (between the two dotted lines) and during the consequent static phase ( $t(23) = 3.9, P < 0.0007$ , comparison for the entire response period). Responses to the static body postures presented directly following opening of the shutter (conditions 3 and 4) are shown in Fig. 3b, revealing a much greater effectiveness of the static articulated posture compared to the non-articulated control postures ( $t(23) = 5.6, P < 0.00001$ ). This difference in response strength lasted for at least 2 s. In Fig. 3c, the responses to the effective articulated action (condition 1) and to the static articulated end-posture (condition 3) are superimposed to facilitate direct comparison. Despite the striking resemblance during the early response, the responses to the articulated actions were significantly larger than to the static articulated postures ( $t(23) = 3.2, P < 0.004$ , comparison for the entire response period).

## 4. Discussion

This study showed that populations of STSa cells are specifically responsive to articulated body actions, while similar non-articulated body actions do not excite these cells. For instance, a cell may respond to a head rotating over the left shoulder toward the subject, but not to a similar rotation

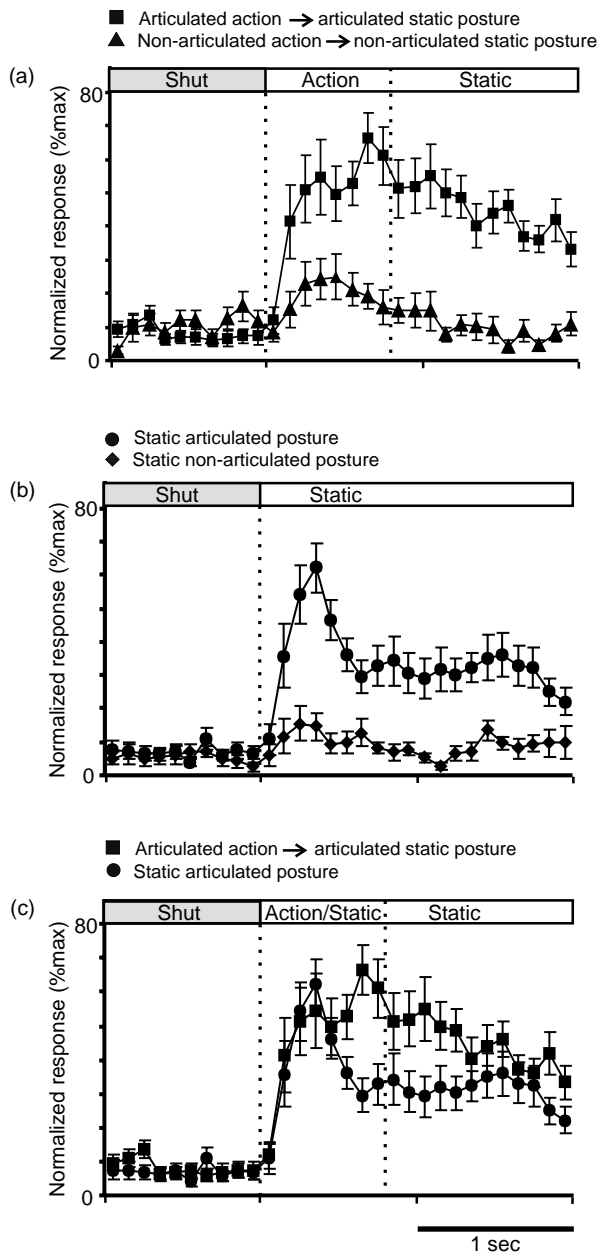


Fig. 3. Population responses to the four main stimulus types. Normalized cell responses (expressed as a percentage of the maximal response of the cell) are given for 100 ms time-bins ( $\pm$ S.E.). (a) Population responses to the specific articulated action the cell was tuned to, ending in the static articulated posture (■), and responses to similar but non-articulated whole body actions, ending in the static non-articulated posture (▲); (b) population responses to articulated static postures (●) and to similar but non-articulated static postures (◆), both without preceding actions; (c) the population responses to the static articulated posture (●) are superimposed on the responses to the articulated actions (■). Mean (non-normalized) response to the articulated action was 29 spikes/s, and to the static articulated end-posture 20 spikes/s.

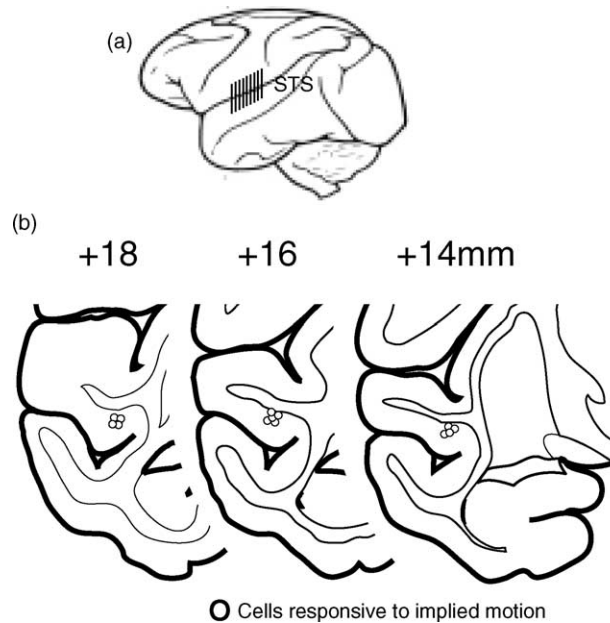


Fig. 4. Histological reconstruction of cell locations. (a) Left side view of the macaque brain. Cells were recorded in the banks of the STSa, between 11 and 19 mm anterior to the inter-aural plane (indicated by vertical bars); (b) reconstruction of coronal sections of the left hemisphere taken at 18, 16 and 14 mm anterior to the inter-aural plane. Each section represents a 2 mm thick slice. All cells for which reconstructions were made were located in the upper bank of STSa. Thick line, cortical surface; thin lines, edge of grey matter.

of the head *and* trunk toward the subject. Furthermore, the static articulated end-posture of the action presented in isolation never failed to evoke a response significantly greater than the response to other non-articulated postures (e.g. the starting posture of the action).

#### 4.1. Do the STSa cells code for implied biological motion?

The findings give rise to the intriguing possibility that the STSa cells code for a particular articulated action both when actually presented and when implied in a still image. Coding for implied biological motion has not yet been studied at the cellular level. Cells in STSa selectively responsive to views of particular body postures have been studied previously (Perrett et al., 1985a; Perrett et al., 1992; Wachsmuth et al., 1994). However, the selectivity described here is novel in several respects. First, cells described previously that were responsive to multiple parts of the body (e.g. head and body) showed tuning for the same view of the head and body (e.g. left profile view of the body and left profile view of the head; Wachsmuth et al., 1994). In the current study, cells showed selectivity for an articulated posture with views of the head turned with respect to the body. Second, we find that cells showed enhancement of response during actions, yet previous studies of cells tuned to articulated postures in the vertical plane (e.g. head raised or lowered) usually did not compare movement and static stimuli or assumed that the cells were sensitive to the particular static body posture per se.

For the population of cells tested here the following findings argue against sole coding for a posture per se, and in

favour of coding for the association of motion and posture: (1) The presentation of the actual articulated action caused a response which was larger than the response to the articulated posture, indicating that the cell was tuned to the action rather than the posture and begun before the final posture was reached. (2) Other, similar, actions, which were not implied in the static articulated posture, did not evoke a response. The cells were selective for one posture, and one articulation which produced that posture. The cells were often insensitive to other movements which could end in the same posture, and were not responsive to articulations starting in the end-posture. (3) The absence of a response to static postures which were very similar except that they did not contain information about the specific implied motion.

#### 4.2. Postural association with preceding and following motion

A static image of an animate object can be associated with movements that are likely to have occurred before the time at which the static image was formed; but the same static image can also be associated with movements that are likely to follow. With a body posture that requires the contraction of one set of muscles, the likely prior motion is that produced by the contraction of those muscles and the likely following motion is that produced by the relaxation of the same muscle set and the constriction of the antagonistic muscle set. Thus, for the static posture depicted in Fig. 1e (back view of the head, left profile view of the body), the likely preceding motion is a rotation of the head from left profile view away from the observer, while the likely following motion is a

rotation of the head to the left profile view (i.e. a return to the ‘resting’ position, with the head and torso aligned and the lateral neck muscles equally contracted/relaxed).

The association with the following motion is as prevalent as with the motion that is likely to have preceded the posture. It is notable that in the present study the association between motion and posture exhibited by cell responses involved movement that brought the body into the static articulated end-point postures. We did not find cell responses tuned to both starting postures and to motions that followed these postures. This asymmetry may reflect a sample bias, or a more fundamental difference in organization.

Systematic bias in association between form and motion has been noted before; two-third of cells selective for whole body motion prefer a combination of body view and direction where the body moves in the direction that the head is pointing (walking forward following the nose); one-third of cells prefer the reverse with the body walking backwards (Oram & Perrett, 1994, 1996; Perrett et al., 1985b). In this case, the association between the forward motion and body view may be that which is most frequently experienced.

#### 4.3. Feedback from STSa onto V5(MT)/MST?

Areas in the ventral object recognition stream (Milner & Goodale, 1995) may play a role in registering the associations between particular forms and associated movements. These higher areas could represent knowledge about the repertoire of behaviour an animate object normally exhibits. Such knowledge would allow an observer to infer whether or not the object was moving at the time at which a fleeting glimpse was made (or a picture was taken), and its most likely previous/next position. These associations revealed in high level visual processing could play a role in influencing activity in processing in V5(MT)/MST through back projections. The activity found in motion sensitive areas V5(MT)/MST when people view static objects with implied motion, but not when the same objects are shown ‘at rest’ (Kourtzi & Kanwisher, 2000; Senior et al., 2000) could reflect such top-down influences. We propose that STSa forms an important source of feedback projections in the case of biological implied motion. Direct and indirect anatomical connections between STSa and V5(MT)/MST exist (e.g. Boussaoud, Ungerleider, & Desimone, 1990).

The human equivalent of the monkey STSa is thought to be the posterior STS region (Karnath, 2001). This area, located just anterior to V5(MT)/MST, typically ‘lights up’ during the viewing of animate actions (see Allison et al., 2000 for a review). If the STS indeed forms a source for the activation of V5(MT)/MST, then one would expect to find activation of STS in fMRI studies of implied motion, in addition to V5(MT)/MST activation. In the study by Kourtzi & Kanwisher (2000), a significant activation of the STS was indeed found in six of the eight subjects in the condition where pictures of athletes with and without implied motion were compared. The study by Senior et al. (2000) did not

find STS activation, but found instead an activation of areas just posterior to V5(MT)/MST (posterior satellites), which these authors attributed to semantic processing of the implied motion images. It is conceivable that activation of the STS region did occur but disappeared after subtraction of the implied and non-implied motion conditions, since both conditions may recruit recognition processing of animate form in STS. The differential activation due to the implied motion per se, might not have been strong enough to survive subtraction. One reason for this could be that the neural populations representing articulated postures (described here) and those representing non-articulated postures (documented elsewhere, Senior et al., 2000; Wachsmuth et al., 1994) are spatially intermingled in STSa. Moreover, the new type of cell we describe sensitive to articulation need not be more numerous than the cell type tuned to non-articulated posture. Furthermore, representations of articulated and non-articulated postures may depend on common building-blocks (i.e. STSa cell populations representing individual parts of the body and insensitive to posture). Each of these explanations would predict that a differential signal to implied motion (or articulated posture) arising from STS could be difficult to detect. Alternatively the STSa could be more relevant for the processing of biological form. The proposed feedback connection from STSa onto V5(MT)/MST fits in with the increasing reports of bi-directional connections between brain areas at different hierarchical stages in the brain (Supèr, Spekreijse, & Lamme, 2001; van Essen & Gallant, 1994).

In conclusion, it may not be usual to think of the visual processing of static form as underlying comprehension of actions. However, this overlooks the ability of individuals to understand momentary postures during an action sequence and to infer how an action was performed. The performance of dexterous manual tasks can easily be specified as a series of static pictures, each demonstrating particular sub-goals or stages in the action sequence. Indeed, the heroes of comic books might not have enjoyed such popularity were it not for our ability to realize what actions are associated with static postures.

#### Acknowledgements

Research was supported by the Human Frontier Science Program and the Wellcome Trust. We thank Bruno Wicker for his contribution to some of the experiments, Mary Latimer for assistance with histology, Robin Edwards for cell location software, and Mike Oram for helpful comments on the manuscript.

#### References

- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences*, 4, 267–278.
- Baker, C. I., Keysers, C., Jellema, T., Wickers, B., & Perrett, D. I. (2000). Coding of spatial position in the superior temporal sulcus



- of the macaque. *Current Psychology Letters: Behaviour, Brain and Cognition*, 1, 71–87.
- Beauchamp, M. S., Lee, K. E., Haxby, J. V., & Martin, A. (2003). Parallel visual motion processing streams for manipulable objects and human movements. *Neuron*, 34, 149–159.
- Bonda, E., Petrides, M., Ostry, D., & Evans, A. (1996). Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *The Journal of Neuroscience*, 16, 3737–3744.
- 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., Shadlen, M. N., Newsome, W. T., & Movshon, J. A. (1992). The analysis of visual motion: a comparison of neuronal and psychophysical performance. *The Journal of Neuroscience*, 12, 4745–4765.
- Britten, K. H., Newsome, W. T., Shadlen, M. N., Celebrini, S., & Movshon, J. A. (1996). A relationship between behavioral choice and the visual responses of neurons in macaque MT. *Visual Neuroscience*, 13, 87–100.
- Bruce, C., Desimone, R., & Gross, C. G. (1981). Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *Journal of Neurophysiology*, 46, 369–384.
- Cohen, M., Kosslyn, S., Breiter, H., DiGorlamo, G., Thompson, W., & Anderson, K. (1995). Changes in cortical activity during mental rotation: a mapping study using fMRI. *Brain*, 119, 89–100.
- David, A. S., & Senior, C. (2000). Implicit motion and the brain. *Trends in Cognitive Sciences*, 4, 293–295.
- Decety, J., Grezes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., Grassi, F., & Fazio, F. (1997). Brain activity during observation of actions: Influence of action content and subject's strategy. *Brain*, 120, 1763–1777.
- Downing, P. E., Yiang, Y. H., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, 293, 2470–2473.
- Dubner, R., & Zeki, S. M. (1971). Response properties and receptive fields of cells in an anatomically-defined region of the superior temporal sulcus in the monkey. *Brain Research*, 35, 528–532.
- Emery, N. J., & Perrett, D. I. (1994). Understanding the intentions of others from visual signals: Neurophysiological evidence. *Current Psychology Cognition*, 13, 683–694.
- Goebel, R., Khorrarn-Sefat, D., Muckli, L., Hacker, H., & Singer, W. (1998). The constructive nature of vision: Direct evidence from functional magnetic resonance imaging studies of apparent motion and motion imagery. *European Journal of Neuroscience*, 10, 1563–1573.
- Grafton, S. T., Arbib, M. A., Fadiga, L., & Rizzolatti, G. (1996). Localization of grasp representations in humans by positron emission tomography. 2. Observation compared with imagination. *Experimental Brain Research*, 112, 103–111.
- Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of Neurophysiology*, 35, 96–111.
- Hoffman, E. A., & Haxby, J. V. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, 3, 80–84.
- Jellema, T., & Perrett, D. I. (2002). Neural coding for visible and hidden objects. *Attention & Performance*, XIX, 356–380.
- Jellema, T., Baker, C. I., Wicker, B., & Perrett, D. I. (2000). Neural representation for the perception of the intentionality of actions. *Brain and Cognition*, 44, 280–302.
- Jellema, T., Baker, C. I., Oram, M. W., Perrett, D. I. (2002). Cell populations in the banks of the superior temporal sulcus of the macaque and imitation. In W. Prinz, & A. Meltzoff (Eds.), *The Imitative Mind* (pp. 267–290). Cambridge: Cambridge University Press.
- Karnath, H. O. (2001). New insights into the functions of the superior temporal cortex. *Nature Reviews Neuroscience*, 2, 568–576.
- Kourtzi, Z., & Kanwisher, N. (2000). Activation in human MT/MST by static images with implied motion. *Journal of Cognitive Neuroscience*, 12, 48–55.
- Maunsell, J. H. R., & van Essen, D. C. (1983). Functional properties of neurons in the middle visual temporal area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. *Journal of Neurophysiology*, 49, 1127–1147.
- Milner, A. D., & Goodale, M. A. (1995). *The Visual Brain in Action*. Oxford: Oxford University Press.
- Oram, M. W., & Perrett, D. I. (1994). Responses of anterior superior temporal polysensory (STPa) neurons to biological motion stimuli. *Journal of Cognitive Neuroscience*, 6, 99–116.
- Oram, M. W., & Perrett, D. I. (1996). Integration of form and motion in the anterior superior temporal polysensory area (STPa) of the macaque monkey. *Journal of Neurophysiology*, 76, 109–1297.
- 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 neurones in the temporal cortex of the macaque monkey: a preliminary report. *Behavioural Brain Research*, 16, 153–170.
- 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. (1989). Frameworks of analysis for the neural representation of animate objects and actions. *Journal of Experimental Biology*, 146, 87–113.
- 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.
- Puce, A., Allison, T., Bentin, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation in humans viewing eye and mouth movements. *The Journal of Neuroscience*, 18, 2188–2199.
- Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., & Fazio, F. (1996). Localization of grasp representations in human by PET. 1. Observation versus execution. *Experimental Brain Research*, 111, 246–252.
- Senior, C., Barnes, J., Giampietro, V., Simmons, A., Bullmore, E. T., Brammer, M., & David, A. S. (2000). The functional neuroanatomy of implicit-motion perception or representational momentum. *Current Biology*, 10, 16–22.
- Shadlen, M. N., Britten, K. H., Newsome, W. T., & Movshon, J. A. (1996). A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *The Journal of Neuroscience*, 16, 1486–1510.
- Smith, A., Greenlee, M., Singh, K., Kraemer, F., & Henning, J. (1998). The processing of first- and second-order motion in human visual cortex assessed by functional magnetic resonance imaging (fMRI). *Journal of Neuroscience*, 18, 3816–3830.
- Supér, H., Spekreijse, H., & Lamme, V. A. F. (2001). Two distinct modes of sensory processing observed in monkey primary visual cortex. *Nature Neuroscience*, 4, 304–310.
- Tootell, R. B. H., Reppas, J. B., Dale, A. M., Look, R. B., Sereno, M. I., Malach, R., Brady, T. J., & Rosen, B. R. (1995). Visual motion after effect in human cortical area MT revealed by functional magnetic resonance imaging. *Nature*, 11, 139–141.
- 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 Behavior* (pp. 549–586). Cambridge, MA: MIT Press.
- van Essen, D. C., & Gallant, J. L. (1994). Neural mechanisms of form and motion processing in the primate visual system. *Neuron*, 13, 1–10.
- Wachsmuth, E., Oram, M. W., & Perrett, D. I. (1994). Recognition of objects by their component parts: responses of single units in the temporal cortex of the macaque. *Cerebral Cortex*, 4, 502–522.

Watson, J., Myers, R., Frackowiak, R., Hajnal, J., Woods, R., & Mazziota, J. (1993). Area V5 of the human brain: evidence from a combined study using positron emission tomography and magnetic resonance imaging. *Cerebral Cortex*, 3, 79–94.

Zeki, S., Watson, J. D. G., & Frackowiak, R. S. J. (1993). Going beyond the information given: The relation of illusory visual motion to brain activity. *Proceedings of the Royal Society of London B*, Vol. 252, 215–222.