

Configural Processing of Biological Motion in Human Superior Temporal Sulcus

James C. Thompson,^{1,2} Michele Clarke,¹ Tennille Stewart,¹ and Aina Puce^{1,2,3,4}

¹Center for Advanced Imaging, Robert C. Byrd Health Sciences Center, Departments of ²Radiology and ³Neurobiology and Anatomy, and ⁴Lane Department of Computer Science and Electrical Engineering, West Virginia University, Morgantown, West Virginia 26506-9236

Observers recognize subtle changes in the movements of others with relative ease. However, tracking a walking human is computationally difficult, because the degree of articulation is high and scene changes can temporarily occlude parts of the moving figure. Here, we used functional magnetic resonance imaging to test the hypothesis that the superior temporal sulcus (STS) uses form cues to aid biological movement tracking. The same 10 healthy subjects detected human gait changes in a walking mannequin in two experiments. In experiment 1, we tested the effects of configural change and occlusion. The walking mannequin was presented intact or with the limbs and torso apart in visual space and either unoccluded or occluded by a set of vertical white bars. In experiment 2, the effects of inversion and occlusion were investigated, using an intact walking mannequin. Subjects reliably detected gait changes under all stimulus conditions. The intact walker produced significantly greater activation in the STS, inferior temporal sulcus (ITS), and inferior parietal cortex relative to the apart walker, regardless of occlusion. Interestingly, STS and ITS activation to the upright versus inverted walker was not significantly different. In contrast, superior parietal lobule and parieto-occipital cortex showed greater activation to the apart relative to intact walker. In the absence of an intact body configuration, parietal cortex activity increased to the independent movements of the limbs and torso. Our data suggest that the STS may use a body configuration-based model to process biological movement, thus forming a representation that survives partial occlusion.

Key words: STS; ITS; SPL; biological motion; inversion; occlusion

Introduction

Animals and humans show a remarkable ability to recognize movements of other biological forms in complex visual scenes, and this recognition is facilitated by specialized brain mechanisms that detect biological motion (Blake, 1993; Oram and Perrett, 1994; Bonda et al., 1996; Puce et al., 1998). Despite our behavioral ability to recognize subtle changes in another's movements, tracking an articulated human walker is computationally complex because there may be 20–40 degrees of freedom for the motion (Deutscher et al., 2000; Aggarwal and Park, 2004). One way to reduce this complexity is to track each moving body part independently. However, this strategy cannot handle conditions of partial occlusion well, because the number of visible moving parts changes (Aggarwal and Park, 2004). This strategy may also be inefficient, because independently tracking multiple moving objects in complex visual scenes is attentionally demanding (Pylyshyn and Storm, 1988; Yantis, 1992). Applying a coarse model of the body could potentially resolve these segmentation ambiguities and assist if tracking breaks down (Marr and Vaina,

1982; Park and Aggarwal, 2002). Using a model or template of body configuration (the relative spatial arrangement of head, limbs, and torso) to track human movement could more effectively segregate the human form from a patterned background or partially occluding foreground.

Human psychophysical studies demonstrate that configural information improves recognition of biological movement (Chatterjee et al., 1996; Shiffrar et al., 1997; Pavlova and Sokolov, 2000; Beintema and Lappe, 2002). Importantly, representing configural information from biological motion stimuli internally may override ambiguous or conflicting environmental cues (Shiffrar et al., 1997; Bulthoff et al., 1998). Hence, using some model of body configuration to aid recognition might solve this perceptual problem. Neuroimaging studies indicate that lateral temporal cortex, in particular the superior temporal sulcus (STS) region, responds preferentially to biological movement (for review, see Puce and Perrett, 2003). However, there is little understanding of the processes that the STS might use to perform this activity. The STS clearly prefers the articulated motion characteristic of human motion to unarticulated motion (Beauchamp et al., 2002). However, these authors observed a clear STS response to unarticulated human movement, suggesting that other factors also contribute to the response.

Here, we investigated the effect of varying body configuration on the response of the brain to detecting gait changes in an articulated dynamic human walker using functional magnetic resonance imaging (fMRI). We varied the spatial configuration of an

Received May 26, 2005; revised Aug. 4, 2005; accepted Aug. 24, 2005.

M.C. and T.S. were funded by the National Science Foundation's Research Environments for Undergraduates Program (Principal Investigator: Dr. Frances Vanscoy, Lane Department of Computer Science and Engineering at West Virginia University). We thank Mary Pettit for editorial assistance.

Correspondence should be addressed to Dr. A. Puce, Center for Advanced Imaging, P.O. Box 9236, West Virginia University, Morgantown, WV 26506-9236. E-mail: apuce@hsc.wvu.edu.

DOI:10.1523/JNEUROSCI.2129-05.2005

Copyright © 2005 Society for Neuroscience 0270-6474/05/259059-08\$15.00/0

animated figure while maintaining identical rotational motion of the component parts. Subjects viewed intact and fragmented versions of the figure, presented either unoccluded or with partial occlusion. A second experiment examined effects of figure inversion and occlusion. We hypothesized that the STS would use configural information to process biological movement, responding reliably even when the moving and intact figure was partially occluded. In the absence of an intact body configuration, we expected that viewing fragmented limbs and torso would preferentially engage other brain regions involved in tracking multiple moving objects, such as the posterior parietal cortex (PPC) (Culham et al., 1998, Jovicich et al., 2001).

Materials and Methods

Subjects

Ten neurologically normal subjects (five males; aged 24–36 years) consented to participate in a study approved by the Institutional Review Board of West Virginia University. All subjects had either normal or corrected-to-normal vision.

Experimental design

Subjects detected gait changes in an animated mannequin whose limbs and body were shown in a walking movement with no net spatial translation on a white background (Fig. 1*A,B*). We chose a mannequin as a walker instead of a point-light or real human figure, because conventional surface form cues such as hair, clothing, and gender would be absent, yet sufficient shape information would be maintained when the configuration was fragmented so that each part could still be easily identified. Previous studies have demonstrated that point-light figures broken into subcomponents are difficult to identify (Pinto and Shiffrar, 1999). Additionally, we reasoned that fragmentation of the mannequin would not be viewed as potentially implausible or as emotionally distressing as viewing a moving, but fragmented, human body. Experimental conditions were presented using a block design to maximize statistical power. In experiment 1, we investigated the effect of intactness and occlusion (Fig. 1*A*) on brain activation. Subjects viewed a tan-colored upright walking mannequin on a white background in the following four conditions in a two-by-two design: (1) unoccluded and intact, (2) occluded and intact, (3) unoccluded and apart, and (4) occluded and apart.

Experiment 2 examined the effects of stimulus inversion and occlusion (Fig. 1*B*) on the intact mannequin. We reasoned that if the effects of intactness were because of the intact mannequin being more familiar than the apart mannequin, we should also see an effect of inversion. Subjects again viewed a tan-colored walking mannequin in the following four conditions in a two-by-two design: (1) unoccluded and upright, (2) occluded and upright, (3) unoccluded and inverted, and (4) occluded and inverted.

In both experiments, each stimulus condition was presented in a block of 30 s duration alternating with a 16 s baseline condition consisting of a uniform white background between each block. Subjects kept their gaze on a red central fixation cross at all times during the entire imaging run, and the displayed stimuli were centered on the fixation cross. Two blocks of each stimulus type were presented in each imaging run, and there were two imaging runs per experiment. The presentation order for the two experiments was randomized across subjects. Each imaging run had between three and seven gait transitions. The interval between each gait transition was drawn randomly from a modified gamma distribution with a mean of 4.5 s (ranging from 2 to 14 s) and filled with gait repetitions. Finally, the first gait transition in each block was randomized across each block.

Stimulus generation

The intact, apart, and inverted versions of the mannequin were created and adjusted using Poser 4.0 (Curious Labs, Santa Cruz, CA). The intact mannequin covered a total area of $2 \times 3^\circ$ of horizontal and vertical visual angle, respectively. The combined elements of the apart mannequin covered $3 \times 3^\circ$ of horizontal and vertical visual angle, respectively; however, each body part size was identical to the intact condition. The intact

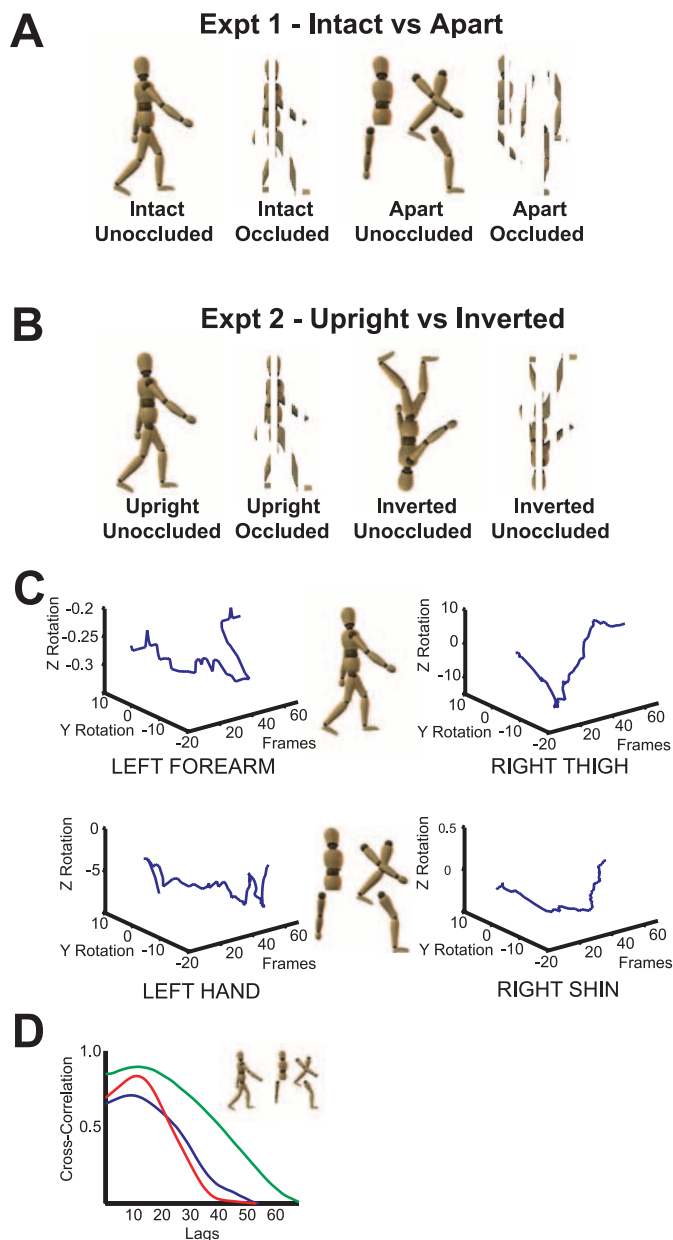


Figure 1. Stimulus conditions. **A**, Experiment 1: intact unoccluded, intact occluded, apart unoccluded, and apart occluded stimuli. **B**, Experiment 2: upright unoccluded, upright occluded, inverted unoccluded, and inverted occluded stimuli. **C**, Examples of *Y* (anteroposterior) and *Z* (inferior–superior) rotation of one gait from two parent segments and their children as a function of video frame (parent, left arm; children, left forearm, left hand; parent, right leg; children, right thigh, right shin). Note that the same *Y*, *Z* (and *X*, not shown) rotation for each segment was used in the intact and apart conditions. **D**, Cross-correlations of *Y* and *Z* rotation from the left hand and left forearm (blue), right thigh (red), and right shin (green) *Y* rotation as a function of frame. The cross-correlation between *Y* (or *Z*) rotations of different segments was identical across intact and apart conditions. Expt, Experiment.

mannequin and the montage of components making up the apart mannequin were presented in the center of the display. Occluders consisted of vertical white bars isoluminant with the white background and were 0.2° thick with 0.2° spacing. There were a total of 19 occluders in each visual display. To animate the mannequin, we used motion files that allowed for the separation of overall position of each body part from the rotation in space around a fixed central point of each part during the gait cycle. Biovision hierarchy (.BVH) files of two gait styles from Mantis Motion Productions (Draper, UT) were obtained, depicting respective gait patterns of a male and female walker. The two gait styles were easily distin-

guishable and differed in terms of arm and leg swing and upper and mid-body rotation. We adjusted the raw .BVH files so that the two gait patterns would have similar speed and step size. The .BVH files consisted of the following: (1) a header containing a definition of a skeleton hierarchy (i.e., whole body) and the offset of each hierarchy parent segment and its children that defined the initial pose and data containing the three-dimensional (3D) position of the hierarchy as a whole, in X (medial–lateral), Y (anterior–posterior), and Z (inferior–superior) coordinates and (2) the 3D rotation of each hierarchy segment across time (in X , Y , and Z coordinates) that defined the gait.

For the present study, the skeleton hierarchy contained five parent segments (torso, left arm, right arm, left leg, right leg), each with segment children [e.g., parent (torso); children (head, neck, abdomen, chest)]. The .BVH files were first edited to make a figure walk in a straight path for two full cycles with identical step size. As our figure walked with no net translation, the position coordinates were constant through the entire gait cycle. Next, we created smooth transitions in Matlab 6.1 (MathWorks, Natick, MA) from one gait to the next (gait transition) and from the end of the gait to the start of the same gait (gait repetition) using the following spatiotemporal morphing algorithm based on that used by Giese and Lappe (2002): $M_{AB} = ((A_i + i)1 - (i \times R)) + B_1(i \times R)$, where M_{AB} = morph from gait A to gait B in X , Y , Z rotation coordinates from each hierarchy segment; $A = X, Y, Z$ rotation coordinates from each hierarchy segment of the initial gait; $B = X, Y, Z$ rotation coordinates from each hierarchy segment of the target gait; R = frame rate of the .BVH file (in seconds); and i = frames to be morphed, where i is $1 \dots n$ frames to be adjusted.

Hence, the morph from gait A to gait B (M_{AB}) consisted of a linear combination of the two gaits that incrementally increased the weight from one gait to the next. A three-point moving average filter was applied to smooth the motion trajectories. These modified .BVH files with the two gait repetitions and two transitions were then applied in Poser 4.0 to the intact, apart, and inverted versions of the mannequin. As a result, the X , Y , and Z rotation of each parent segment was the same in the intact (Fig. 1C) and apart (Fig. 1D) and upright and inverted conditions. Hence, the relationship between the children of each segment was the same in the various conditions, and only the overall position of each parent segment differed. This procedure maintained higher-order temporal correlations between the motion of different segments (Figs. 1E, F). Poser figures were then exported as animations [audio video interleave (.AVI) format] using a duration of 2.3 s and a frame rate of 30 frames/s. A foreground scene with occluders was also generated in Poser 4.0, enabling occluded versions of the mannequin to be created. Brief animations showing a gait transition for a subset of the stimuli can be viewed on-line as supplemental material (available at www.jneurosci.org).

The .AVI files with gait repetitions and transitions were imported into Adobe Premiere 6.5 (Adobe Systems, San Jose, CA). Successive .AVI files were concatenated, resulting in 30 s stimulus blocks in which the mannequin walked with a gait that smoothly changed from one style to another at random intervals during the imaging run. Stimulus blocks were alternated with a 16 s white diffuse screen. A red central fixation cross was superimposed on the center of the display for the duration of the entire stimulus run. The animations for each imaging run were exported from Adobe Premiere as .AVI files and converted into moving pictures experts group (MPEG3) format with ArcSoft Showbiz DVD software (Hewlett-Packard, Palo Alto, CA) and outputted to digital video disk (DVD).

Stimulus delivery

Stimuli were projected through the control room window of the MRI scanner onto a screen at the subject's feet using a video projector (LT10 DLP; NEC, Tokyo, Japan) connected to a DVD player (DVP-NS325; Sony, Tokyo, Japan). The subject lay supine in the MRI scanner and viewed the display through a mirror mounted on the quadrature head coil. Subjects were instructed to respond by button press when they detected a gait transition but to maintain their gaze on a fixation cross at all times. In experiment 1, the numbers of random gait transitions for each condition were the following: unoccluded intact, 20; occluded intact, 19; unoccluded apart, 22; occluded apart, 20. For experiment 2, the

numbers of transitions were the following: unoccluded upright, 23; occluded upright, 19; unoccluded inverted, 20; occluded apart, 18.

Data acquisition

Functional images. We acquired 188 volumes per run of blood oxygen level-dependent (BOLD) activity from posterior temporal and parietal cortex using a gradient echo spiral-in–out sequence (Glover and Law, 2001) on a 3 tesla Horizon LX MRI scanner (General Electric Medical Systems, Milwaukee, WI). Functional images used in subsequent analyses consisted of average images of spiral-in and spiral-out trajectories, reconstructed off-line using routines written in C (courtesy of G. Glover, Stanford University, Stanford, CA) and running under Linux (version 7.1; RedHat, Raleigh, NC). Seventeen coronal slices from the posterior commissure to the occipital pole (3 mm thick with 1 mm gap) were acquired with the following parameters: matrix, 128×128 ; field of view (FOV), 240 mm (in plane resolution, 1.875 mm); bandwidth, 125; and echo time/repetition time, 25/2000 ms.

Anatomical images. We acquired a part-brain T1-weighted anatomical volume with identical slice prescription to functional images and a whole-brain high-resolution anatomical spoiled gradient-recalled acquisition in a steady state (SPGR) volume ($1.2 \times 0.9375 \times 0.9375$ mm; FOV, 240; matrix, 256×256 ; 248 slices with 50% overlap).

Data analysis

Data analysis was performed using Oxford Centre for Functional Magnetic Resonance Imaging of the Brain Linear Image Registration Tool (FLIRT; FMRIB Software Library, Oxford, UK) and SPM2 (Functional Imaging Laboratory, University College London, London, UK), Caret 5.3 (Van Essen laboratory, Washington University, St. Louis, MO), and Matlab. Coregistration of the partial-brain coronal T1 to the high-resolution SPGR was more optimal using FLIRT relative to SPM2, as assessed by identification of anatomical landmarks on visual inspection. All subsequent processing steps were performed using SPM2. Coronal functional images were coregistered to the coronal T1 template and corrected for motion and slice-timing differences. The SPGR image was normalized to the Montreal Neurological Institute template, and the parameters determined for this normalization were applied to functional images. Voxels from the SPGR and functional images were resampled to 2 mm^3 . Finally, the functional images were smoothed with a Gaussian kernel with a full width at half maximum of 8 mm.

Given our hypotheses, we acquired coronal slices from temporal and parietal cortex; hence, only voxels from these regions featured in our analyses. The WFUPickAtlas software (ANSIR Core; Wake Forest School of Medicine, Wake Forest, NC) was used to create a temporal and parietal cortex mask that was applied to the normalized functional images from the two experiments. For both experiments, regressors for each of the experimental conditions were created by modeling the BOLD response to each stimulus condition as a box car with an "ON" cycle equal to the length of each block, convolved with a hemodynamic response function and its temporal derivative. Included in the model were six motion covariates (three translation and three rotation parameters) determined from the motion correction and a constant term to account for drift. A high-pass temporal filter of 140 s was also applied to the data to remove low-frequency drifts in the MR signal. Parameter estimates were created for each subject, and contrasts comparing the experimental conditions to fixation, as well as the main effects and interactions from experiments 1 and 2, were created. The significance of each of these contrasts was examined by using one-sample t tests at the group level in a random effects analysis.

We first compared the experimental conditions from both experiment 1 and experiment 2 relative to their respective fixation baselines, using a significance threshold of $p < 0.005$ (uncorrected) and a >4 contiguous voxel cluster as inclusion criteria, because our hypotheses for possible differences between conditions were based on specific cortical regions (lateral temporal cortex and posterior parietal cortex). We then used regions that showed a significant positive response in experiment 1 or 2 as volumes of interest in which to compare the various conditions. Comparisons between the experimental conditions were thresholded for significance at $p < 0.005$ (uncorrected) and a >4 contiguous voxel cluster.

Results

Behavioral data

Accuracy of detecting gait changes was similar across conditions in both experiments (means for experiment 1: unoccluded and intact, 99.0%; occluded and intact, 99.8%; unoccluded and apart, 97.0%; occluded and apart, 94.7%; all $p > 0.1$; means for experiment 2: unoccluded and upright, 96.2%; occluded and upright, 96.9%; unoccluded and inverted, 94.2%; occluded and inverted, 94.7%; all $p > 0.1$). Response times were not recorded with our experimental set-up.

fMRI data

Experiment 1

Brain regions showing significant effects of intactness or occlusion are listed in Table 1. Greater activation to the intact relative to apart walker, regardless of occluders, was seen in bilateral STS ($p < 0.001$), bilateral Sylvian cortex/inferior parietal cortex (right, $p < 0.005$; left, $p < 0.001$), right inferior temporal sulcus (ITS), extending into the inferior temporal gyrus ($p < 0.001$), right intraparietal sulcus ($p < 0.005$), and left inferior parietal lobule ($p < 0.005$) (see Figs. 2, 5). Conversely, there was significantly greater activation to the apart relative to intact walker in bilateral superior parietal lobule (SPL) (left, $p < 0.001$; right, $p < 0.005$), right parieto-occipital cortex ($p < 0.005$), and left middle temporal gyrus (MTG) ($p < 0.005$) (see Figs. 3, 5). Responses were greater ($p < 0.005$) in left fusiform gyrus, ITS, and right SPL to the unoccluded than occluded walker, regardless of whether the walker was intact or apart (data not shown). In contrast, a greater response ($p < 0.005$) to occluded relative to unoccluded walker was observed in dorsal structures only (i.e., left Sylvian and inferior medial parietal cortex; data not shown). A significant interaction between occlusion and intactness was observed in right STS and right posterior ITS (both $p < 0.005$) (Figs. 4, 5). The right STS location of this interaction was lateral to the activation of the main effect of intactness. The posterior ITS focus was consistent with previous reports of a lateral temporal region selective for the visual presentation of human bodies, called the extrastriate body area (EBA) (Downing et al., 2001). Results indicated a greater response to the intact versus apart walker in the unoccluded condition only in both the right STS and ITS ($p < 0.005$).

Experiment 2

Brain regions showing significant effects of inversion or occlusion are listed in Table 2. Significantly greater activation to the upright relative to inverted walker was observed in the right parahippocampal gyrus ($p < 0.001$), right SPL ($p < 0.001$), bilateral posterior ITS (right, $p < 0.001$; left, $p < 0.005$), and left fusiform gyrus ($p < 0.005$). Greater responses to the unoccluded relative to the occluded walker were observed in left inferior parietal lobule ($p < 0.001$), left middle occipital gyrus ($p < 0.001$), left ITS ($p < 0.005$), and right fusiform gyrus ($p < 0.005$). We observed no significantly greater responses to the inverted versus upright walker or to the occluded versus unoccluded walker, and the inversion-by-occlusion interaction was also not significant. No region showing a greater response to the main effect of intact

Table 1. Experiment 1: main effects of occlusion and intactness and their interactions

Anatomical location as a function of contrast	Talairach coordinates			Volume of activated cortex (mm ³)
	x	y	z	
Intact > apart				
Superior temporal sulcus	−64	−52	14	152
Sylvian/IPL	−40	−52	22	392
Sylvian/IPL	−40	−40	32	64
Inferior temporal sulcus	46	−50	−6	880
Superior temporal sulcus	44	−40	16	704
Inferior parietal lobule	34	−34	34	64
Intraparietal sulcus	42	−38	36	40
Apart > intact				
Middle temporal gyrus	−54	−58	6	64
Parieto-occipital cortex	−26	−80	22	56
Superior parietal lobule	−22	−74	36	600
Intraparietal sulcus	30	−66	56	80
Postcentral gyrus	34	−52	64	248
Unoccluded > occluded				
Fusiform gyrus	−38	−52	−22	288
Superior parietal lobule	−30	−48	62	176
Inferior temporal sulcus	38	−64	−6	152
Occluded > unoccluded				
Sylvian cortex	−36	−48	22	104
Inferior parietal cortex	−20	−62	34	104
Intactness × occlusion				
Inferior temporal sulcus	−38	−66	10	64
Superior temporal sulcus	64	40	10	136

Anatomical locations of peak activation and corresponding Talairach coordinates are listed. IPL, Inferior parietal lobule.

versus apart walker showed a significant difference between the upright versus inverted walker (all $p > 0.05$). The right posterior ITS region that showed a significant interaction between intactness and occlusion in experiment 1 showed a greater response to the upright versus inverted walker (see Figs. 4, 5).

Discussion

We investigated how body configuration contributes to the processing of biological movement in temporal and parietal cortex using fMRI. Supporting our original hypothesis, the STS, in addition to other lateral temporal and inferior parietal regions, showed greater responses to a moving walker with intact body configuration relative to a moving form in which limbs were fragmented from the torso. The movement of individual body parts was always the same in our intact and apart stimulus conditions, and the spatial configuration of the body parts was manipulated. The preferential response to the intact walker was observed despite the presence of partial occlusion. Our data suggest that the STS uses a body shape-based model to process biological movement, identifying a neural correlate of the body-shape model proposed in computational vision and theoretical and psychophysical studies (Marr and Vaina, 1982; Beintema and Lappe, 2002; Aggarwal and Park, 2004). The apart–intact dimension showed a dorsal–ventral activation gradient in parietotemporal cortex; brain regions showing preferentially greater responses to the apart walker were generally more dorsal to those sensitive to the intact walker (i.e., the bilateral SPL, right parieto-occipital cortex, and left MTG).

Selectivity for intact biological movement in lateral temporal cortex

Many neuroimaging studies have shown the sensitivity of the STS and surrounding lateral temporal regions to biological movement (Bonda et al., 1996; Puce et al., 1998; Grezes et al., 2001; Grossman and Blake, 2001, 2002; Vaina et al., 2001; Beauchamp

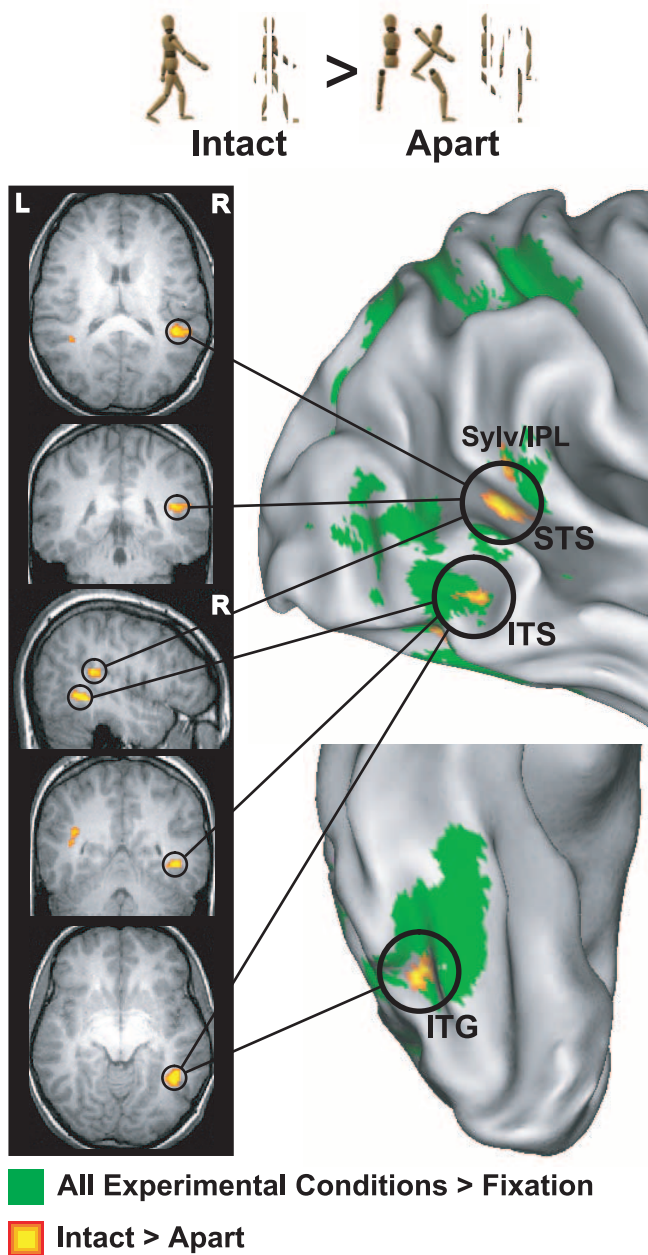


Figure 2. Experiment 1: intact versus apart walker. Group analysis activation data overlaid on a right lateral and a right ventral view of the inflated Colin brain atlas anatomical image (right) and a single subject's normalized SPGR anatomical image (insets at left). Warm colors depict voxels where activation was significantly greater for intact versus apart. Green depicts voxels that activated significantly to all experimental conditions from experiments 1 or 2 versus fixation baseline. ITG, Inferior temporal gyrus; Sylv/IPL, Sylvian/inferior parietal lobule; L, left; R, right.

et al., 2002; Pelphrey et al., 2003). STS preference for biological movement contrasts with the neighboring MTG region, which is sensitive to the motion of tools and other forms of mechanistic motion (Beauchamp et al., 2002; Blakemore et al., 2003; Pelphrey et al., 2003). Beauchamp and colleagues (2002) demonstrated that the STS prefers the articulated motion characteristic of human movement relative to more rigid motion characteristic of tools or to human figures moving in a rigid, artificial manner. Our data support and extend these findings by suggesting that rather than tracking the motion of individual body parts independently, the STS tracks a moving figure holistically, using the presence of intact body configuration. Crucially, our stimulus

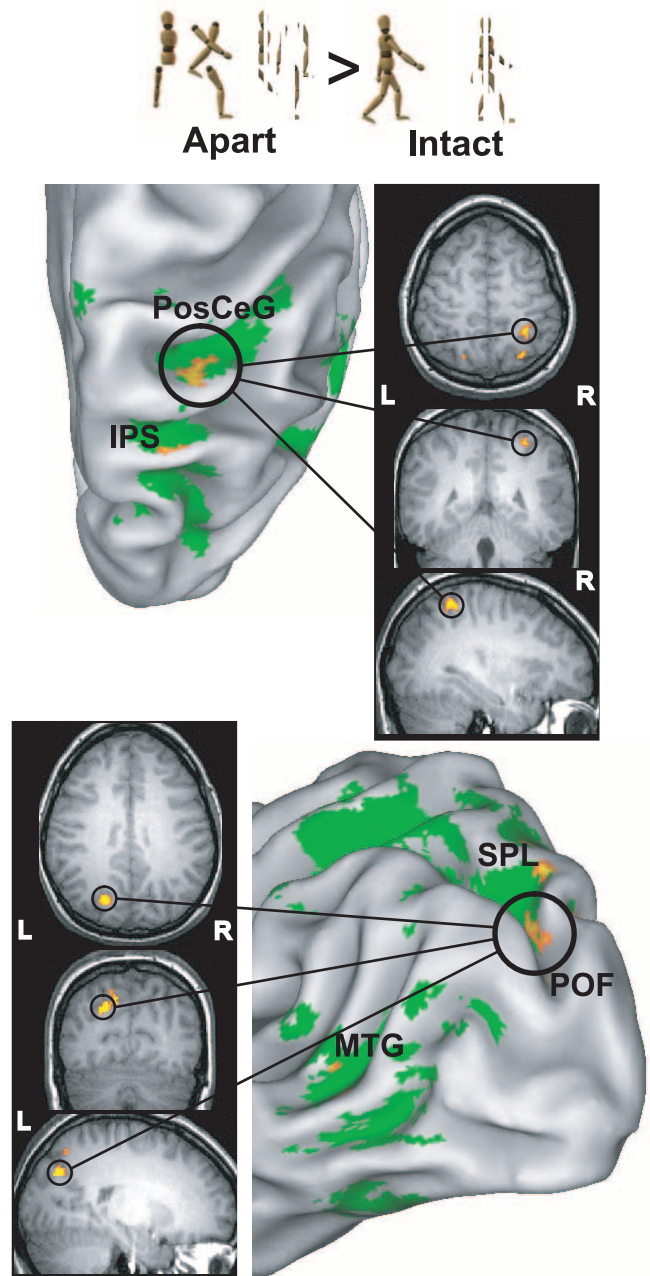


Figure 3. Experiment 1: apart versus intact walker. Group analysis activation data overlaid on right dorsal (top) and left ventrolateral (bottom) hemispheric views of the inflated Colin brain atlas anatomical images and a single subject's normalized SPGR anatomical image (insets) are shown. Warm colors depict voxels where activation was significantly greater for apart versus intact. Green depicts voxels that activated significantly to all experimental conditions from experiments 1 or 2 versus fixation baseline. PosCeG, Posterior central gyrus; IPS, intraparietal sulcus; POF, parieto-occipital fissure; L, left; R, right.

maintained higher-order temporal correlations between the movements of individual body parts even when the figure was fragmented. Hence, the STS does not process biological movement simply by monitoring the temporal correlation between independently moving objects but instead requires these moving objects to be configured in a manner consistent with a body shape.

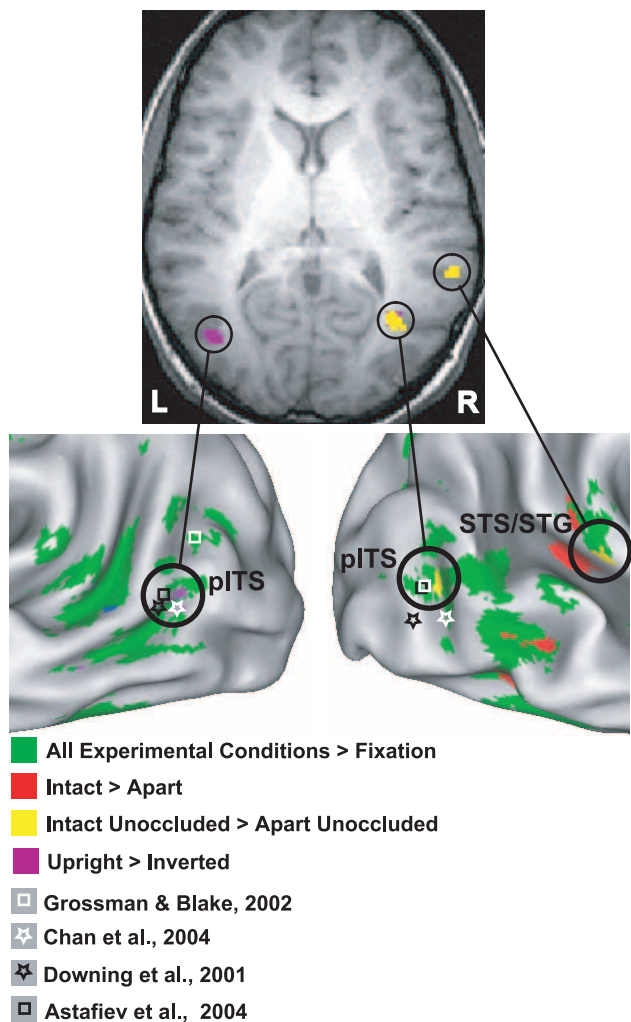


Figure 4. Overlap in effects between experiments 1 and 2. Group analysis activation data overlaid on a single subject's normalized SPGR images (top) and inflated Colin brain atlas anatomical images of the left (bottom left) and right (bottom right) lateral hemispheres. Green indicates significant activation to all experimental conditions from experiments 1 or 2 versus fixation. Red indicates a significantly greater response to intact versus apart, regardless of occlusion from experiment 1. Yellow indicates interaction between intactness and occlusion, from experiment 1. Purple indicates a significantly greater response to upright versus inverted, regardless of occlusion, from experiment 2. Symbols indicate location of peak activation in EBA in previous studies. STG, Superior temporal gyrus; pITS, inferior temporal sulcus, posterior segment; L, left; R, right.

Greater ITS activation occurred to the intact than apart walker, even when the walker was partially occluded. Others have reported greater activation in inferior temporal cortex to intact point-light biological motion figures than to scrambled point-light figures (Grossman and Blake, 2002; Saygin et al., 2004). The inferior temporal cortex has been implicated in object recognition (Malach et al., 1995; Pietrini et al., 2004), visual imagery of objects (Ishai et al., 2000), and retrieval of visual objects from memory (Wheeler and Buckner, 2003). This region of inferior temporal cortex has also been shown to respond to configuration of object shape rather than the local image features (Kourtzi and Kanwisher, 2001) and may engage in object completion, whereby partially occluded objects are able to be recognized as whole (Lerner et al., 2002). The involvement of object completion in the processing of biological movement is supported by data in temporal lobe lesion patients impaired in the recognition of point-

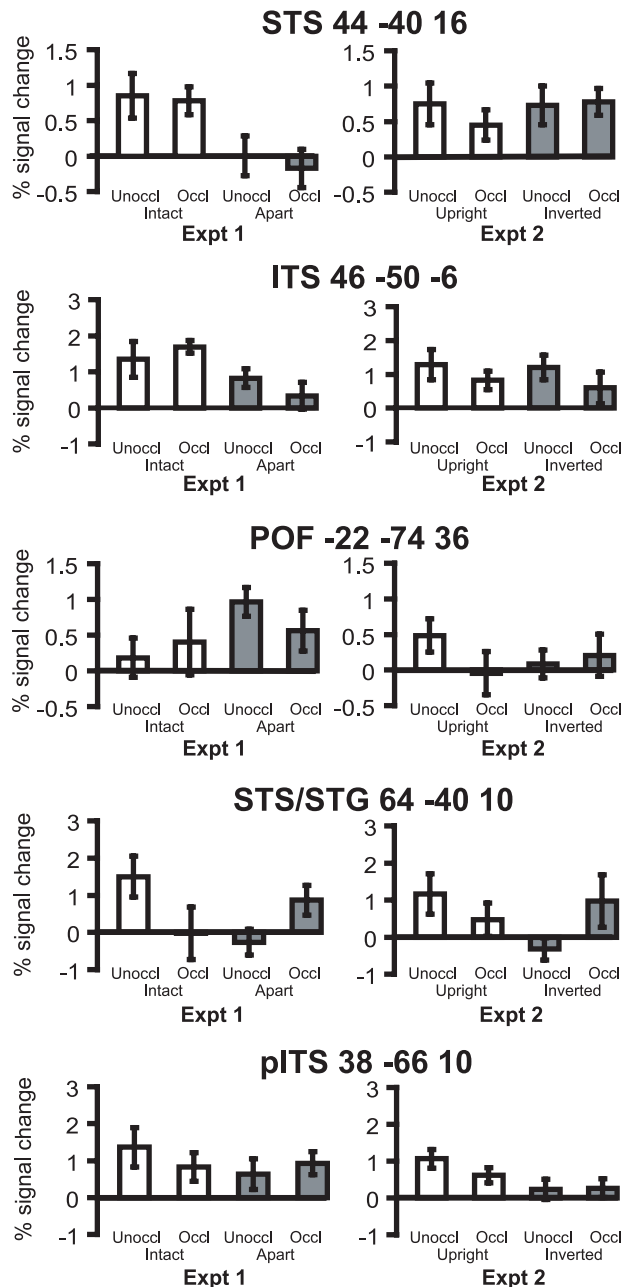


Figure 5. Experiments (Expts) 1 and 2: mean \pm 1 SE percentage MR signal change (% signal change) relative to fixation baseline at maximally responding anatomical locations (with corresponding centroids in Talairach coordinates). Data from experiments 1 and 2 are presented in left and right columns, respectively. White bars represent intact (Expt 1) or upright (Expt 2) conditions, gray-filled bars represent apart (Expt 1) or inverted (Expt 2) conditions. The activations for the right STS and right ITS were presented originally in Figure 2, left POF in Figure 3, and right STS/STG and right pITS in Figure 4. Unoccl, Unoccluded; Occl, occluded; POF, parieto-occipital fissure; STG, superior temporal gyrus; pITS, inferior temporal sulcus, posterior segment.

light biological motion as well as recognition of degraded, incomplete static objects (Vaina and Gross, 2004).

Our findings illustrate how form cues, such as body configuration, influence the way motion information is processed in the human brain. We suggest that the ITS interacts with the STS to enable biological movement recognition, despite partial occlusion. A high-level, probabilistic representation of body configuration in the ITS may act as a shape prior that is used by the STS together with estimates of articulated or correlated motion to

Table 2. Experiment 2: main effects of occlusion and inversion and their interactions

Anatomical location as a function of contrast	Talairach coordinates			Volume of activated cortex (mm ³)
	x	y	z	
Upright > inverted				
Fusiform gyrus	−48	−48	−14	128
Inferior temporal sulcus	−46	−70	6	96
Parahippocampal gyrus	40	−30	−8	152
Inferior temporal sulcus	36	−62	6	224
Superior parietal lobule	20	−60	64	144
Unoccluded > occluded				
Middle occipital gyrus	−40	−76	−6	128
Inferior temporal gyrus	−44	−70	4	248
Inferior parietal lobule	−46	−46	50	112
Fusiform gyrus	38	−38	−24	96

Anatomical locations of peak activation and corresponding Talairach coordinates are listed.

recover probabilistic details of biological movement. Computer-vision based methods for the recognition of human movement often use body shape-based models to recover motion details (Blake and Isard, 1998; Aggarwal and Park, 2004). Body shape-based models may be particularly useful for recognizing and classifying the movements of a figure partially occluded by scene elements, such as trees or fences, or for identifying holistic changes to a figure, such as those seen with changes to gait or posture. A Bayesian account of the integration of form and motion has been described recently by Sajda and Baek (2004). This model uses the probabilistic representation of shape to modulate velocity estimates of local motion to solve a number of object motion tasks. Although such a model involves a unidirectional influence of shape on motion processing, it is highly likely that shape and motion pathways influence each other in a bidirectional manner (Wallach and O'Connell, 1953; Ungerleider and Mishkin, 1982).

Changing configuration versus inversion

In the present study, we did not observe an inversion effect in most of the regions responding to body configuration changes. Behavioral studies of face processing indicate that face inversion can produce impairments in processing configural information (Rhodes et al., 1993; Farah et al., 1995; Freire et al., 2000; Leder and Bruce, 2000). Similarly, impaired recognition for inverted point-light or line-drawn displays of biological motion (Shiffrar et al., 1997; Pavlova and Sokolov, 2000) or static body positions of computer-generated figures (Reed et al., 2003) has also been documented. However, fMRI studies show variable effects of stimulus inversion in category-specific regions. Unlike in our study, inverted point-light figures have been shown to reduce the STS response relative to an upright point-light figure when subjects passively view stimuli (Grossman and Blake 2001). Viewing inverted faces relative to upright faces in face-selective regions has produced reduced, unchanged, and greater activation in a number of studies (Kanwisher et al., 1998; Aguirre et al., 1999; Haxby et al., 1999; Yovel and Kanwisher, 2004).

Sensitivity to local form features in lateral STS and posterior ITS

In contrast to regions located either more deeply in the STS, or in more anterolateral ITS, regions in right lateral STS and posterior ITS locations appeared to respond more strongly to the intact relative to the apart walker only when the stimulus was unoccluded. Although both foci showed a slightly lower response to the intact walker when occluded, this effect was not significant. A bilateral posterior ITS focus also responded significantly more

strongly to the upright relative to the inverted walker, whereas the lateral STS focus tended to show a nonsignificant trend with this manipulation, and only when the figure was unoccluded. These two foci appear to be more sensitive to changes to local form features of the mannequin. The posterior ITS focus may correspond to a lateral temporal region reported previously to be selective to visual presentation of human bodies, labeled the extrastriate body area (Downing et al., 2001). The EBA responds more strongly to whole bodies than to body parts (Downing et al., 2001) and upright views of lower body parts than inverted views (Chan et al., 2004). Human

motion-sensitive region hMT+ is also in close proximity to this location (Tootell et al., 1995; Dumoulin et al., 2000; Huk et al., 2002). This region has shown some sensitivity to object shape (Kourtzi et al., 2002; Kriegeskorte et al., 2003). Because there was no difference in the motion of individual body parts between the various conditions, we suggest the posterior ITS focus observed in the present study is more likely to correspond to a form processing region such as EBA. However, the functional and anatomical relationships between the EBA and hMT+, and their respective roles in the processing of biological motion, is clearly a subject that will require detailed investigation in future studies.

Tracking of individual moving limbs and torso

Increased responses in bilateral SPL, right parieto-occipital cortex, and left MTG were observed when the limbs and torso were spatially fragmented relative to the intact walker. We reasoned that in the absence of body configuration, subjects would need to track the motion of limbs and torso independently in a manner similar to tracking the motion of independent objects. Subjects were still able to accurately identify gait changes even when the walker was apart, indicating that subjects were able to successfully track the motion pattern of the limbs and torso. Attentional tracking of moving objects has been shown to activate regions of the SPL and parieto-occipital cortex in close proximity to those activated in the present study (Culham et al., 1998; Jovicich et al., 2001). Activation of these regions is not limited to attending to motion, and they may form part of a general attention mechanism in the PPC (Wojciulik and Kanwisher, 1999). However, we would suggest that this is a less efficient means of tracking biological movement and is likely susceptible to competing demands on attention.

Conclusion

The findings of the present study indicate that the processing of biological movement in STS relies on form cues, such as body configuration. Combining form cues with motion information appears to be one method to track the complex, dynamic movements of other living forms. This method provides us with a means of tracking biological movement even in the presence of partial occlusion, a necessity when one is observing another moving in a busy visual scene.

References

- Aggarwal JK, Park S (2004) Human motion: modeling and recognition of actions and interactions. In: Second international symposium on 3D data processing, visualization and transmission, pp 640–647. Washington, DC: IEEE Computer Society.

- Aguirre GK, Singh R, D'Esposito M (1999) Stimulus inversion and the responses of face and object-sensitive cortical areas. *NeuroReport* 10:189–194.
- Astafiev SV, Stanley CM, Shulman GL, Corbetta M (2004) Extrastriate body area in human occipital cortex responds to the performance of motor actions. *Nat Neurosci* 7:542–548.
- Beauchamp MS, Lee KE, Haxby JV, Martin A (2002) Parallel visual motion processing streams for manipulable objects and human movements. *Neuron* 34:149–159.
- Beintema JA, Lappe M (2002) Perception of biological motion without local image motion. *Proc Natl Acad Sci USA* 99:5661–5663.
- Blake A, Isard M (1998) Active contours: the application of techniques from graphics, vision, control theory and statistics to visual tracking of shapes in motion. Berlin: Springer.
- Blake R (1993) Cats perceive biological motion. *Psychol Sci* 4:54–57.
- Blakemore SJ, Boyer P, Pachot-Clouard M, Meltzoff A, Segebarth C, Decety J (2003) The detection of contingency and animacy from simple animations in the human brain. *Cereb Cortex* 13:837–844.
- Bonda E, Petrides M, Ostry D, Evans A (1996) Specific involvement of human parietal systems and the amygdala in the perception of biological motion. *J Neurosci* 16:3737–3744.
- Bulthoff I, Bulthoff H, Sinha P (1998) Top-down influences on stereoscopic depth-perception. *Nat Neurosci* 1:254–257.
- Chan AW, Peelen MV, Downing PE (2004) The effect of viewpoint on body representation in the extrastriate body area. *NeuroReport* 15:2407–2410.
- Chatterjee SH, Freyd JJ, Shiffrar M (1996) Configural processing in the perception of apparent biological motion. *J Exp Psychol Hum Percept Perform* 22:916–929.
- Culham JC, Brandt SA, Cavanagh P, Kanwisher NG, Dale AM, Tootell RB (1998) Cortical fMRI activation produced by attentive tracking of moving targets. *J Neurophysiol* 80:2657–2670.
- Deutscher J, Blake A, Reid I (2000) Articulated body motion capture by annealed particle filtering. *Comput Vis Pattern Recognit* 2:126–133.
- Downing PE, Jiang Y, Shuman M, Kanwisher N (2001) A cortical area selective for visual processing of the human body. *Science* 293:2470–2473.
- Dumoulin SO, Bittar RG, Kabani NJ, Baker Jr CL, Le Goualher G, Bruce Pike G, Evans AC (2000) A new anatomical landmark for reliable identification of human area V5/MT: a quantitative analysis of sulcal patterning. *Cereb Cortex* 10:454–463.
- Farah MJ, Tanaka JW, Drain HM (1995) What causes the face inversion effect? *J Exp Psychol Hum Percept Perform* 21:628–634.
- Freire A, Lee K, Symons LA (2000) The face-inversion effect as a deficit in the encoding of configural information: direct evidence. *Perception* 29:159–170.
- Giese MA, Lappe M (2002) Measurement of generalization fields for the recognition of biological motion. *Vision Res* 42:1847–1858.
- Glover GH, Law CS (2001) Spiral-in/out BOLD fMRI for increased SNR and reduced susceptibility artifacts. *Magn Reson Med* 46:515–522.
- Grezes J, Fonlupt P, Bertenthal B, Delon-Martin C, Segebarth C, Decety J (2001) Does perception of biological motion rely on specific brain regions? *NeuroImage* 13:775–785.
- Grossman ED, Blake R (2001) Brain activity evoked by inverted and imagined biological motion. *Vision Res* 41:1475–1482.
- Grossman ED, Blake R (2002) Brain areas active during visual perception of biological motion. *Neuron* 35:1167–1175.
- Haxby JV, Ungerleider LG, Clark VP, Schouten JL, Hoffman EA, Martin A (1999) The effect of face inversion on activity in human neural systems for face and object perception. *Neuron* 22:189–199.
- Huk AC, Dougherty RF, Heeger DJ (2002) Retinotopy and functional subdivision of human areas MT and MST. *J Neurosci* 22:7195–7205.
- Ishai A, Ungerleider LG, Haxby JV (2000) Distributed neural systems for the generation of visual images. *Neuron* 28:979–990.
- Jovicich J, Peters RJ, Koch C, Braun J, Chang L, Ernst T (2001) Brain areas specific for attentional load in a motion-tracking task. *J Cogn Neurosci* 13:1048–1058.
- Kanwisher N, Tong F, Nakayama K (1998) The effect of face inversion on the human fusiform face area. *Cognition* 68:B1–B11.
- Kourtzi Z, Kanwisher N (2001) Representation of perceived object shape by the human lateral occipital complex. *Science* 293:1506–1509.
- Kourtzi Z, Bulthoff HH, Erb M, Grodd W (2002) Object-selective responses in the human motion area MT/MST. *Nat Neurosci* 5:17–18.
- Kriegeskorte N, Sorger B, Naumer M, Schwarzbach J, van den Boogert E, Hussy W, Goebel R (2003) Human cortical object recognition from a visual motion flowfield. *J Neurosci* 23:1451–1463.
- Leder H, Bruce V (2000) When inverted faces are recognized: the role of configural information in face recognition. *Q J Exp Psychol A* 53:513–536.
- Lerner Y, Hendler T, Malach R (2002) Object-completion effects in the human lateral occipital complex. *Cereb Cortex* 12:163–177.
- Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB (1995) Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci USA* 92:8135–8139.
- Marr D, Vaina L (1982) Representation and recognition of the movements of shapes. *Proc R Soc Lond B Biol Sci* 214:501–524.
- Oram MW, Perrett DI (1994) Responses of anterior superior temporal polysensory (STPa) neurons to “biological motion” stimuli. *J Cogn Neurosci* 6:99–116.
- Park S, Aggarwal JK (2002) Segmentation and tracking of interacting human body parts under occlusion and shadowing. In: *Proceedings of the workshop on motion and video computing*, pp 105–111. IEEE Computer Society: Washington, DC.
- Pavlova M, Sokolov A (2000) Orientation specificity in biological motion perception. *Percept Psychophys* 62:889–899.
- Pelphrey KA, Mitchell TV, McKeown MJ, Goldstein J, Allison T, McCarthy G (2003) Brain activity evoked by the perception of human walking: controlling for meaningful coherent motion. *J Neurosci* 23:6819–6825.
- Pietrini P, Furey ML, Ricciardi E, Gobbi ML, Wu WH, Cohen L, Guazzelli M, Haxby JV (2004) Beyond sensory images: object-based representation in the human ventral pathway. *Proc Natl Acad Sci USA* 101:5658–5663.
- Pinto J, Shiffrar M (1999) Subconfigurations of the human form in the perception of biological motion displays. *Acta Psychol* 102:293–318.
- Puce A, Perrett D (2003) Electrophysiology and brain imaging of biological motion. *Philos Trans R Soc Lond B Biol Sci* 358:435–445.
- Puce A, Allison T, Bentin S, Gore JC, McCarthy G (1998) Temporal cortex activation in humans viewing eye and mouth movements. *J Neurosci* 18:2188–2199.
- Pylshyn ZW, Storm RW (1988) Tracking multiple independent targets: evidence for a parallel tracking mechanism. *Spat Vis* 3:179–197.
- Reed CL, Stone VE, Bozova S, Tanaka J (2003) The body-inversion effect. *Psychol Sci* 14:302–308.
- Rhodes G, Brake S, Atkinson AP (1993) What's lost in inverted faces? *Cognition* 47:25–57.
- Sajda P, Baek K (2004) Integration of form and motion within a generative model of visual cortex. *Neural Netw* 17:809–821.
- Saygin AP, Wilson SM, Hagler Jr DJ, Bates E, Sereno MI (2004) Point-light biological motion perception activates human premotor cortex. *J Neurosci* 24:6181–6188.
- Shiffrar M, Lichter L, Heptulla Chatterjee S (1997) The perception of biological motion across apertures. *Percept Psychophys* 59:51–59.
- Tootell RB, Reppas JB, Dale AM, Look RB, Sereno MI, Malach R, Brady TJ, Rosen BR (1995) Visual motion aftereffect in human cortical area MT revealed by functional magnetic resonance imaging. *Nature* 375:139–141.
- Ungerleider LG, Mishkin M (1982) Two cortical visual systems. In: *Analysis of visual behavior* (Dingle J, ed), pp 549–586. Cambridge, MA: MIT.
- Vaina LM, Gross CG (2004) Perceptual deficits in patients with impaired recognition of biological motion after temporal lobe lesions. *Proc Natl Acad Sci USA* 101:16947–16951.
- Vaina LM, Solomon J, Chowdhury S, Sinha P, Belliveau JW (2001) Functional neuroanatomy of biological motion perception in humans. *Proc Natl Acad Sci USA* 98:11656–11661.
- Wallach H, O'Connell DN (1953) The kinetic depth effect. *J Expt Psychol* 45:205–247.
- Wheeler ME, Buckner RL (2003) Functional dissociation among components of remembering: control, perceived oldness, and content. *J Neurosci* 23:3869–3880.
- Wojciulik E, Kanwisher N (1999) The generality of parietal involvement in visual attention. *Neuron* 23:747–764.
- Yantis S (1992) Multielement visual tracking: attention and perceptual organization. *Cognit Psychol* 24:295–340.
- Yovel G, Kanwisher N (2004) Face perception: domain specific, not process specific. *Neuron* 44:889–898.