

Joakim Vinberg and Kalanit Grill-Spector

J Neurophysiol 99:1380-1393, 2008. First published Jan 2, 2008; doi:10.1152/jn.01223.2007

You might find this additional information useful...

Supplemental material for this article can be found at:

<http://jn.physiology.org/cgi/content/full/01223.2007/DC1>

This article cites 64 articles, 31 of which you can access free at:

<http://jn.physiology.org/cgi/content/full/99/3/1380#BIBL>

Updated information and services including high-resolution figures, can be found at:

<http://jn.physiology.org/cgi/content/full/99/3/1380>

Additional material and information about *Journal of Neurophysiology* can be found at:

<http://www.the-aps.org/publications/jn>

This information is current as of March 20, 2008 .

Representation of Shapes, Edges, and Surfaces Across Multiple Cues in the Human Visual Cortex

Joakim Vinberg¹ and Kalanit Grill-Spector^{1,2}

¹Psychology Department and ²Neurosciences Institute, Stanford University, Stanford, California

Submitted 2 November 2007; accepted in final form 26 December 2007

Vinberg J, Grill-Spector K. Representation of shapes, edges, and surfaces across multiple cues in the human visual cortex. *J Neurophysiol* 99: 1380–1393, 2008. First published January 2, 2008; doi:10.1152/jn.01223.2007. The lateral occipital complex (LOC) responds preferentially to objects compared with random stimuli or textures independent of the visual cue. However, it is unknown whether the LOC (or other cortical regions) are involved in the processing of edges or global surfaces without shape information. Here, we examined processing of 1) global shape, 2) disconnected edges without a global shape, and 3) global surfaces without edges versus random stimuli across motion and stereo cues. The LOC responded more strongly to global shapes than to edges, surfaces, or random stimuli, for both motion and stereo cues. However, its responses to local edges or global surfaces were not different from random stimuli. This suggests that the LOC processes shapes, not edges or surfaces. LOC also responded more strongly to objects than to holes with the same shape, suggesting sensitivity to border ownership. V7 responded more strongly to edges than to surfaces or random stimuli for both motion and stereo cues, whereas V3a and V4 preferred motion edges. Finally, a region in the caudal intraparietal sulcus (cIPS) responded more strongly to both stereo versus motion and to stereo surfaces versus random stereo (but not to motion surfaces vs. random motion). Thus we found evidence for cue-specific responses to surfaces in the cIPS, both cue-specific and cue-independent responses to edges in intermediate visual areas, and shape-selective responses across multiple cues in the LOC. Overall, these data suggest that integration of visual information across multiple cues is mainly achieved at the level of shape and underscore LOC's role in shape computations.

INTRODUCTION

The human visual cortex is composed of numerous visual areas whose retinotopy and functional characteristics have been extensively studied (for a review see Grill-Spector and Malach 2004). However, the organizational principles that underlie the specialization of visual areas remain mysterious. In particular, it is unclear to what extent processing of primary visual cues such as motion (Albright et al. 1984; Castelo-Branco et al. 2002; Heeger et al. 1999; Huk and Heeger 2002; Newsome et al. 1989; Sack et al. 2006; Tootell et al. 1995; Zeki 2004), stereo (Sakata et al. 1999; Taira et al. 2000; Tsao et al. 2003), or luminance are segregated into distinct cortical regions, or to what extent visual areas integrate different visual cues toward performing a specific visual task, such as object recognition (Goodale and Milner 1992; Mishkin et al. 1983). Previous research has provided evidence for both segregation and convergence of visual cue processing in the human and

primate visual cortex (DeAngelis et al. 1998; Ffytche et al. 1995; Grill-Spector and Malach 2004; Palanca and DeAngelis 2003; Seidemann et al. 1999; Shikata et al. 2001; Tsutsui et al. 2002, 2005).

Several studies provide evidence for cue-invariant processing of objects in the human and primate ventral stream. Neurons in macaque inferotemporal cortex respond more strongly to shapes than to simple features and their shape selectivity is maintained across different visual cues (Sary et al. 1993; Tanaka et al. 2001; Vogels and Orban 1996). The human lateral occipital complex (LOC; Grill-Spector et al. 2001; Malach et al. 1995), a region thought to be involved in object recognition (Grill-Spector et al. 2000, 2001), responds more strongly to objects defined from either luminance, texture, illusory contours, motion, stereo, or color than to random stimuli or uniform textures generated from the same primary cues (Appelbaum et al. 2006; Gilaie-Dotan et al. 2002; Grill-Spector et al. 1998; Kastner et al. 2000; Kourtzi and Kanwisher 2000; Mendola et al. 1999). These studies provide evidence for convergence of low-level visual cues in generating object-selective responses in the LOC.

Visual processing en route to object recognition must be able to extract shape¹ independently of the visual cues that define it. The recognition process likely entails a sequence of computations across visual cortex, starting from local computations in early visual cortex related to low-level properties of the visual stimulus, such as disparity, motion, or orientation, conveying little sense of the global object shape, then proceeding to more global computations in higher levels of the hierarchy of visual processing. However, it is unknown whether the LOC (or other visual regions) are involved in the processing of intermediate visual information such as edge or surface information that may be extracted during processing en route to object recognition (Marr 1982; Nakayama 1995; Nakayama et al. 1989). In particular, Nakayama and colleagues (Nakayama 1995; Nakayama et al. 1989) suggested that surfaces are represented in intermediate visual areas of the visual processing stream. This surface representation may provide a means to segment the visual input into relevant regions, which in turn may enhance the efficiency of recognition processes because they

¹ Our working definition of an *edge* is a discontinuity in the image. Edges can be real or illusory. Our working definition of a *shape* is a distinct region enclosed by an edge. Shapes, by definition, are bounded by an edge. However, edges do not always define a shape, e.g., disconnected edges (Fig. 1C) do not define a global shape.

Address for reprint requests and other correspondence: K. Grill-Spector, Department of Psychology, Stanford University, Stanford, CA 94305 (E-mail: kalanit@psych.stanford.edu).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

would operate only on a subset of meaningful surfaces, rather than the entire visual input (Nakayama et al. 1995).

Previous studies have examined the processing of edge information or surface information en route to recognition only from luminance cues. Several studies have shown that the LOC activates more strongly for global shapes defined by luminance edges than for misaligned, incomplete, or scrambled luminance edges (Altmann et al. 2003; Doniger et al. 2000; Lerner et al. 2002, 2004), suggesting that LOC responds to global shapes defined by luminance edges rather than to local luminance edges. However, it is unknown 1) whether this pattern of results is specific for luminance-defined stimuli or extends to other cues and 2) whether the LOC also processes local edge information as compared with random stimuli with no edges whatsoever.

Others have suggested that the LOC processes global surface information because its response to enclosed regions defined by Kanizsa inducers was higher than its response to rotated inducers, even when the illusory contours were not perceived (Stanley and Rubin 2003). However, it is unknown 1) whether this pattern of response is specific to luminance-defined stimuli and 2) whether the LOC responds to global surfaces without shape or edge information.

To address these questions, we investigated the involvement of visual regions in processing global shapes, global surfaces, or edges by examining differential cortical activation across these stimuli. Critically, we examined whether preferential responses for shapes, surfaces, or edges were cue specific, or occurred across both motion and stereo cues. We conducted two functional magnetic resonance imaging (fMRI) experiments in which subjects viewed dot stimuli generated from one visual cue (motion or stereo) that created either 1) a percept of a global shape (Fig. 1, *A* and *B*), or 2) disconnected edges that did not form a global shape or a global surface (Fig. 1*C*), or 3) two global surfaces without edges (Fig. 1*D*), or 4) were completely random (Fig. 1*E*). Random stimuli had the same relative disparity (or motion) range as that of other conditions, but contained no coherent structure, edges, or global surfaces.

We reasoned that regions involved in processing global surfaces would respond more strongly to surfaces than to random stimuli that are otherwise identical except that they contain no coherent structure (Fig. 1, *D* vs. *E*). The surfaces

condition created a percept of two semitransparent global surfaces in two depth planes (generated via disparity or motion parallax) without any edges. We posited that regions involved in processing edges would respond more strongly to stimuli containing edges (Fig. 1*C*) than either to surfaces or to random stimuli. The edges condition contained disconnected motion or stereo-edges generated from scrambling object contours without forming a global shape. Finally, we reasoned that regions involved in global shape processing would activate more strongly to shape (Fig. 1, *A* and *B*) than to edges, surfaces, or random stimuli. To better characterize the involvement of LOC in shape processing we created two types of shapes—objects and holes—that had the same shape and contour but differed in their border ownership (Fig. 1, *A* vs. *B*).

All stimuli were made of random dots of the same density, contained no luminance edges (except for the screen boundary that was present during the entire experiment, including the blank baseline periods), and were controlled for low-level properties (see METHODS). We chose to generate stimuli from motion and stereo cues because 1) they allow control of low-level properties across conditions, 2) perception of edges is possible from these cues alone, 3) both motion parallax and disparity provide sufficient information for perceiving surfaces and their relative depth, and 4) processing of motion and stereo cues is thought to involve distinct cortical regions, which allowed us to separate cue processing from content processing.

METHODS

Subjects

Thirteen subjects (seven female, six male; age range 21–36 yr) participated in the study. All subjects had normal or corrected-to-normal vision and intact stereoscopic vision. Subjects provided informed consent to participate in the experiment, which was approved by the Stanford Internal Review Board on Human Subjects Research. Subjects viewed all stimuli through red/cyan glasses.

fMRI acquisition

Subjects were scanned on a 3.0T GE scanner located at the Lucas Imaging Center, Stanford University. We used a receive-only surface coil that covered the occipital lobe, the posterior parietal lobe, and most of the temporal lobe. We acquired T1-weighted in-plane ana-

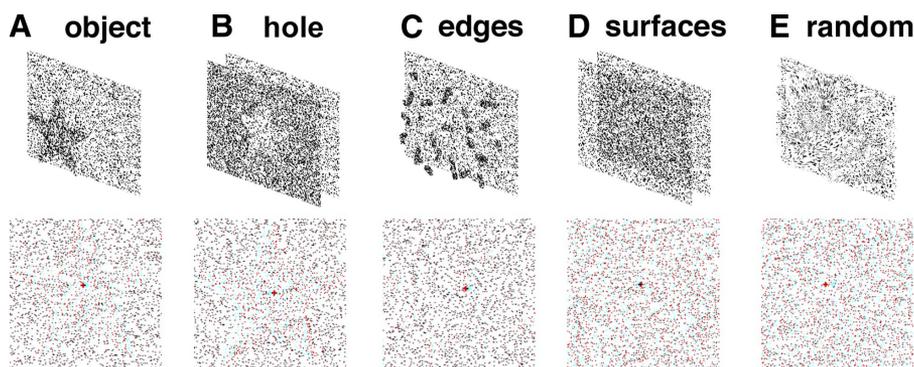


FIG. 1. Experimental conditions. *Top*: schematic illustration of experimental conditions; darker regions indicate front surfaces; stimuli had no luminance edges (except for the screen border that was present during the entire experiment including blank (fixation) baseline blocks). *Bottom*: illustration of stereo conditions (use red/cyan glasses, red on left eye). Noise during printing may incorrectly render some stereo-surfaces to appear nonflat. To view full screen stereo and motion stimuli watch online Supplemental Movies S1 and S2. *A*: object on the front surface in front of a flat background plane. *B*: hole on the front surface in front of a flat background. *C*: disconnected edges in front of a flat background. *D*: 2 semitransparent flat surfaces at different depths. *E*: random stimuli with no coherent structure, edges, global surfaces, or global shape. Random stimuli had the same relative disparity or depth range as that of other conditions.

tomical images at the same prescription as the functional data using an SPGR (spoiled gradient recalled acquisition in steady-state) pulse sequence [repetition time (TR) = 2,000 ms, echo time (TE) = 1.9 ms, flip angle = 15°, field of view (FOV) = 20 cm, bandwidth = 15.63]. We collected functional data from 32 slices oriented perpendicular to the calcarine, with a voxel size of $3.125 \times 3.125 \times 3$ mm, and a TR = 2,000 ms using a spiral sequence (Glover 1999). We acquired whole brain high-resolution anatomical images using a head coil in a separate session for 11 of 13 subjects.

Subjects' behavioral data were collected via a scanner-compatible response box. One subject's behavioral data during the scan were incompletely recorded and thus excluded from analysis.

General experimental design

Stimuli were presented in a block design in which 16 s blocks of visual stimuli alternated with 8 s blank periods that consisted of a gray background with a centrally presented fixation. Each trial in a block lasted 2 s. The blank (fixation) baseline was identical across motion and stereo experiments. The only luminance edge was the screen border, which was present in all conditions including the blank condition (see online Supplemental Movies S1 and S2).² In each of the stereo and motion experiments, subjects viewed six blocks of each experimental condition, totaling 48 unique trials per condition. Shapes were identical across objects and holes and across stereo and motion experiments, and included silhouettes of animals, cars, common objects, geometric forms, and abstract shapes. Object and hole sizes were $11.5 \pm 1.8^\circ$. Since there is no significant difference in the level of LOC activation across nameable and nonnameable shapes (Supplemental Fig. S1), using both types of shapes enabled us to relate our findings to previous experiments that used both nameable (Grill-Spector et al. 1998; Kourtzi and Kanwisher 2000) and nonnameable shapes (Kourtzi and Kanwisher 2000; Stanley and Rubin 2003).

Stereo experiments

Stimuli were stationary random-dot stereograms with 8% dot density (Fig. 1 and online Supplemental Movie S1). All conditions used the same range of relative disparity (± 0.24 to 0.37°). There were no luminance edges in any of the conditions. Each stereogram contained two depth-defined surfaces, except the random condition in which dots were assigned all possible disparities in a random assignment. Each trial in a block (lasting 2 s) contained a different pairing of disparities. Conditions: 1) Stereo-object (Fig. 1A): superposition of two random-dot patterns, a shaped flat region at a closer disparity and a farther background plane. 2) Stereo-hole (Fig. 1B): a front random-dot surface with a shaped hole through which a flat random-dot surface at the farther disparity could be seen. 3) Stereo-edges (Fig. 1C): disconnected stereo-edges presented at a single depth in front of a flat background presented at a farther disparity. Edges were generated by: *i*) scrambling shape outlines, *ii*) expanding the edge 3 pixels on either side, *iii*) filling this region with random dots of density 8%, and *iv*) placing these edges at a nearer plane than the background random-dot plane. The width of edges was 0.36° . There were no luminance edges and stereo-edges did not form a global shape. The edge stimuli had the same total contour length as that of objects and holes and edges, as in previous studies (Altmann et al. 2003; Lerner et al. 2002, 2004). 4) Stereo surfaces (Fig. 1D): two semitransparent planes filling the entire display, at two different depths with no edges, created by the superposition of two random-dot patterns (front and back planes) that were assigned different disparities. Semitransparent surfaces were created by the following procedure: *i*) we generated two different random-dot images, *ii*) each of the patterns was assigned a different disparity and colored red or cyan, and *iii*) we superimposed these two images into one image by taking the minimum color value

in each pixel across the two images. That is, if a pixel was colored in at least one image, it was assigned the minimum color; otherwise, it was assigned to white. 5) Random (Fig. 1E): dots were assigned a random disparity and appeared as a three-dimensional (3D) cloud of dots with no structure.

In each object or hole block, the objects were presented centrally in half of the trials and displaced horizontally by $6.7 \pm 0.7^\circ$ in the other half of the trials. Thus the centrally presented fixation was on the front surface of the object (or hole) for half of the trials in the block and on the farther (background) surface for the other half of the trials. The object or hole was displaced to equalize responses in the behavioral task (where subjects judged whether the fixation was on the front or back surface) in object and hole trials to the other experimental conditions.

TASK. Subjects were asked to fixate and judge in each trial whether the stereoscopically coded fixation point overlaid the nearer or farther surface; during blank trials subjects pressed a key when the fixation cross blinked. Fixation was assigned to the near surface in half the trials and the far surface in the other half, in the random condition the fixation was in front of, or behind, the dot cloud. This task requires parsing of the image across all experimental conditions and is orthogonal to the contrasts of interest. Performance was similar across conditions (see *Behavioral responses during motion and stereo experiments*).

Stereo *experiments 1* and *2* were identical except that *experiment 1* did not include the edges condition. Nine subjects participated in the first experiment, eight subjects in the second, and four in both experiments.

Motion experiments

Dot density was 8% and dot lifetime was 100%. Each 2 s trial in a block was composed of 20 frames, each 100 ms long, to create a percept of visual motion. There were no disparity and no luminance edges. Surfaces, edges, object, and hole conditions contained a stationary background plane made of random dots and a front region (made of another random-dot pattern) that moved coherently to the right in half the trials and to the left in the other half of the trials (online Supplemental Movie S2). Speeds ($3\text{--}3.5^\circ/\text{s}$) and motion directions were randomly assigned across trials. Conditions: 1) Motion-object: a shaped occluding random-dot surface moved across a stationary background of random dots. 2) Motion-hole: an occluding random-dot surface with a shaped hole moved over the stationary background surface made of random dots. 3) Motion-edges: object outlines were scrambled to create disconnected edges. These edges were filled with random dots (same as stereo experiment) and moved in the same direction and speed across a stationary random-dot background. There were no luminance edges and motion-edges did not form a global shape. 4) Motion-surfaces: a front semitransparent plane composed of random dots moved across a stationary background plane composed of random dots; motion parallax provided information of which plane was in front. 5) Random: dots moved randomly and motion speed was similar to that of other conditions. Here in each 2 s trial we generated 20 different random images, which created the percept of random motion.

TASK. Subjects were asked to fixate and judge the direction of movement of the moving surface (left or right); during blank trials subjects pressed a key when the fixation cross blinked. This task required image parsing in all experimental conditions and was orthogonal to the contrasts of interest. Performance was similar across conditions and across stereo and motion experiments (see RESULTS).

Motion *experiments 1* and *2* were identical except that *experiment 1* did not include the edges condition. The same nine subjects participated in motion and stereo *experiment 1*. The same eight subjects participated in motion and stereo *experiment 2*.

² The online version of this article contains supplemental data.

Localizers and retinotopy

All subjects participated in an independent block-design LOC localizer scan in which they viewed gray-level photographs of cars, faces, animals, and novel sculptures, and scrambled versions of these images. Subjects viewed six blocks of each condition and covertly named the stimuli as cars, faces, animals, sculptures and scrambled while fixating. This experiment was used to define LOC regions of interest (ROIs) (see Fig. 3).

All subjects participated in a middle temporal (MT) localizer in which they viewed blocks of 16 s low-contrast expanding and contracting rings that alternated with 16 s low-contrast stationary rings. Subjects viewed 12 blocks of each condition. MT was defined in each subject as voxels that responded more strongly to moving stimuli compared with stationary stimuli located at the posterior end of the inferotemporal sulcus (ITS), thresholded at $P < 10^{-5}$, uncorrected (Tootell et al. 1995; see Fig. 2).

Eleven (of 13) subjects participated in meridian mapping and eccentricity scans (Grill-Spector and Malach 2004) that were used to delineate retinotopic areas that are defined by visual field maps: V1, V2, V3, V3a, V4, and V7 (see Fig. 2).

fMRI data analysis

Data were preprocessed and coregistered to a high-resolution anatomical volume using BrainVoyager 2000 (Brain Innovation, Maas-tricht, The Netherlands). All subsequent analyses were performed in BrainVoyager QX and in-house code written in Matlab (The Math-Works, Natick, MA).

Functional data were temporally filtered with a linear trend removal and high-pass filtering with a cutoff of three cycles per run. We registered the functional data for each subject to their high-resolution anatomical volume. Subjects' high-resolution anatomies were transformed to Talairach coordinates, segmented to gray and white matter, and their cortical surface was reconstructed.

Voxel-by-voxel analysis

Voxel-based analyses used a general linear model (GLM) implemented in BrainVoyager QX. Statistical maps were generated from GLM analyses that were done on a subject-by-subject basis.

ROI analyses

ROIs were defined for each subject on their inflated cortex (to restrict ROIs to the gray matter) and then projected to the 3D volume

from which time-course data were extracted. Figure 2 shows the location of all ROIs on one representative subject's inflated cortical surface. Talairach coordinates and ROI volumes are presented in Table 1.

We extracted the average time course of each ROI and the percentage signal change was calculated relative to the fixation baseline. Amplitudes were calculated by averaging the percentage signal change 4–18 s following the onset of a block across the six repetitions of each block.

LOC ROIs

LOC was defined as lateral and ventral occipital regions that activated more strongly to gray-level images of objects (animals, cars, and abstract sculptures) than to scrambled versions of these objects (Fig. 3A). This contrast activates a large cortical extent in lateral and ventral occipitotemporal cortex. Typically, the LOC is separated into two regions—LO and posterior fusiform gyrus/occipitotemporal sulcus (pFUS/OTS)—using a combination of anatomical and functional criteria (Grill-Spector et al. 1999, 2000; Sayres and Grill-Spector 2006).

LO. A region in the lateral occipital cortex, posterior and adjacent to MT, that responded significantly more strongly to gray-level objects than to scrambled objects ($P < 0.0001$, voxel level, uncorrected). We excluded voxels that overlapped MT. We divided LO into 2 sub-ROIs (Fig. 2). The first ROI [marked “(a)” in Figs. 2 and 3D] included the dorsal aspect of LO and overlapped a retinotopic representation. This ROI extended into cortex defined as LO2 (Larsson and Heeger 2006). However, we did not find a consistent upper visual meridian and therefore we could not define the ROI using the same retinotopic criteria as those defined by Larsson and colleagues. The more ventral part of LO that extended beyond the retinotopic region was taken as a second ROI marked “(b)” in Figs. 2 and 3D. In 2 of 13 subjects without a retinotopic map, LO was subdivided using anatomical cues.

Our definition of LO is unlikely to overlap with that of kinetic occipital (KO) (Dupont et al. 1997; Tyler et al. 2006; Van Oostende et al. 1997) because the functional definition and spatial location of these regions is different. KO refers to a region between V3a and MT that responds more strongly to kinetic gratings than to uniform motion (Dupont et al. 1997; Van Oostende et al. 1997) and to stereo gratings versus a single plane (Tyler et al. 2006), but it does not activate to luminance-defined gratings (Van Oostende et al. 1997). KO overlaps with LO1, a retinotopic region adjacent and inferior to V3a (Larsson and Heeger 2006), but LO is ventral and lateral to LO1.

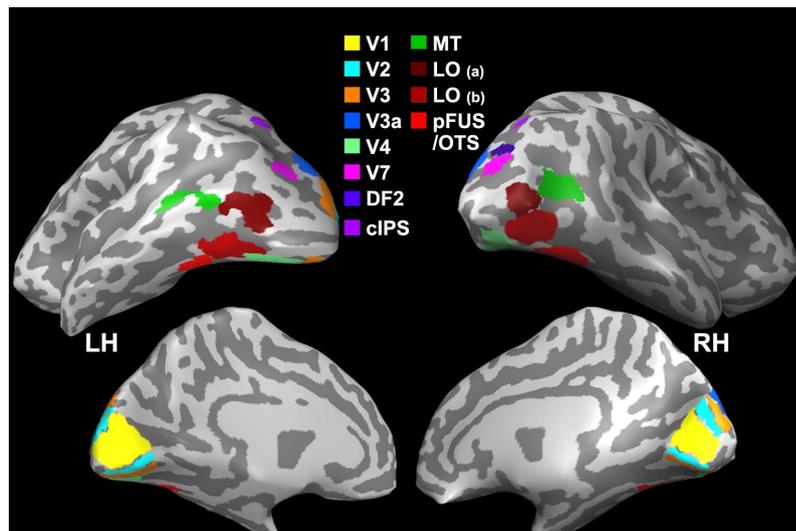


FIG. 2. Regions of interest (ROIs) on an inflated cortical surface of a representative subject. Lateral occipital complex (LOC) ROIs (LO (a), LO (b), pFUS/OTS) were defined from the gray-level localizer experiment. The middle temporal (MT) area was defined from the motion localizer experiment. Retinotopic visual areas (V1, V2, V3, V3a, V4) were defined from meridian mapping experiments and dorsal ROIs (V7, DF2, cIPS) were defined from the motion and stereo experiments. Left Hemisphere (LH); Right Hemisphere (RH).

TABLE 1. *Talairach coordinates and size of ROIs*

ROI	Left Hemisphere				Right Hemisphere			
	X, mm	Y, mm	Z, mm	Size, mm ³	X, mm	Y, mm	Z, mm	Size, mm ³
<i>A. LOC ROIs</i>								
LO								
(a)	-39 ± 3	-79 ± 4	-3 ± 5	1,800 ± 573	41 ± 4	-74 ± 9	-6 ± 5	1,647 ± 748
(b)	-43 ± 5	-68 ± 4	-9 ± 5	1,409 ± 551	41 ± 3	-69 ± 4	-11 ± 4	1,446 ± 393
pFUS/OTS	-40 ± 4	-56 ± 5	-13 ± 4	901 ± 419	38 ± 4	-56 ± 7	-15 ± 5	902 ± 554
<i>B. Dorsal ROIs</i>								
V7	-27 ± 6	-79 ± 3	15 ± 4	624 ± 278	27 ± 5	-77 ± 4	17 ± 6	684 ± 481
DF2	-19 ± 4	-79 ± 4	19 ± 4	744 ± 111	26 ± 8	-78 ± 5	25 ± 9	638 ± 489
cIPS	-21 ± 6	-64 ± 10	42 ± 7	475 ± 180	23 ± 7	-69 ± 13	40 ± 8	576 ± 222

Values are means ± SD across subjects.

pFUS/OTS. A ventral region overlapping the fusiform gyrus and occipitotemporal sulcus that responded more strongly to gray-level objects than to scrambled objects ($P < 0.0001$, voxel level, uncorrected; Figs. 2 and 3D). In previous studies we referred to this region as pFUS (Grill-Spector et al. 1999, 2000; Sayres and Grill-Spector 2006).

Dorsal ROIs

We defined three ROIs along the intraparietal sulcus (IPS) on a subject-by-subject basis from responses during motion and stereo experiments. 1) The most posterior ROI (Figs. 2 and Fig. 6C) was defined as a region in the posterior IPS that responded more strongly to both motion-edges versus random motion and to stereo-edges versus random stereo (each contrast, $P < 0.01$, voxel level). This ROI was adjacent to V3a and was located at the posterior end of the IPS. Superposition of this ROI on retinotopic maps showed that this region overlapped with the retinotopic representation of V7 (Press et al. 2001). Because this ROI overlapped with the visual field map of V7 defined from retinotopic scans, we will use the term V7 when referring to this ROI to link between our results and the existing literature (e.g., Tsao et al. 2003; Tyler et al. 2006). 2) A region in

caudal IPS (cIPS) was defined based on its higher responses to stereo-edges than to random-stereograms ($P < 0.01$, voxel level, uncorrected; Figs. 2 and 6C) and was detected in all eight subjects who participated in *experiment 2*. In five subjects the same region responded more strongly for motion-edges than to random motion ($P < 0.01$, voxel level, uncorrected). We labeled this ROI cIPS, based on its definition in previous studies (Tsao et al. 2003). 3) In six subjects we found an IPS region anterior to V7 and posterior to cIPS that responded more strongly to stereo-surfaces than to random-surfaces ($P < 0.01$, voxel level, uncorrected) and we analyzed responses of this region separately (DF2, Figs. 2 and 6D).

Measuring the overlap between activation to objects from motion and objects from stereo

To quantify the overlap between regions responding to objects from motion and objects from stereo, we ran a voxel-by-voxel GLM detecting voxels that responded more strongly for stereo objects than for random stereo ($P < 0.005$, voxel level, Fig. 3B) and a second GLM detecting voxels that responded more strongly for motion objects than for random motion ($P < 0.005$, voxel level, Fig. 3C). Activations overlapped the LOC, extending anteriorly along the lateral bank of the brain and to the

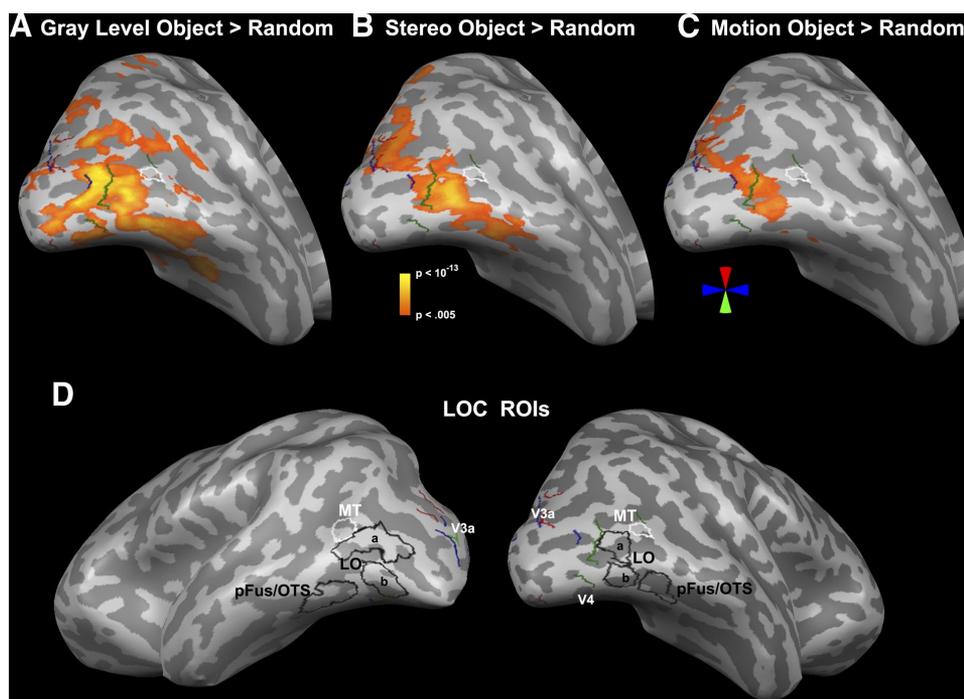


FIG. 3. Selective responses to objects across multiple visual cues in the LOC. *Top*: statistical maps of objects > random ($P < 0.005$, voxel level) from gray-level (A), motion (B), and stereo experiments (C), on the inflated left hemisphere of a representative subject. LOC activated more to objects than random from all 3 cues. Visual meridians are represented by the red, blue, and green lines: red, upper; blue, horizontal; green, lower. White contour: MT. *D*: LOC ROIs defined from higher activation to objects than random in the gray-level localizer experiment.

posterior portion of the fusiform gyrus and occipitotemporal sulcus. For each subject we quantified the size of the region that activated for stereo-objects, motion-objects, and the overlapping region that was activated by both contrasts. We excluded voxels identified as MT. We measured in each subject the fraction of the overlapping region relative to the region activated by stereo-objects or motion-objects and calculated the mean and SD across subjects.

Psychophysics of the perceived shape of objects and holes

All subjects participated in a psychophysics study conducted outside the scanner approximately 4 months after the fMRI scans using identical stimuli. We measured accuracy and response times to 40 trials of each of the following conditions: stereo-objects, stereo-holes, motion-objects, and motion-holes.

TASK. Subjects were asked to fixate and report whether the fixation was inside or outside the shape on screen. This task tests whether subjects could determine a shaped region in the hole and object conditions. We chose to use this task because it was used previously to evaluate the perceived saliency of shaped regions (Stanley and Rubin 2003) and because many of our stimuli contained abstract shapes that did not allow measurement of explicit object recognition.

RESULTS

Behavioral responses during motion and stereo experiments

Performance during fMRI scans was high and similar across conditions (Table 2). In both stereo *experiments 1* and *2* there were no significant differences in accuracy or response time (RT) across experimental conditions (*t*-test, $P > 0.06$). During the first motion experiment there were no significant differences in accuracy or response times across experimental conditions (*t*-tests, $P > 0.08$) except for higher accuracy and longer RTs for motion-surfaces than for motion-objects (*t*-test, $P < 0.05$). These differences were not significant after Bonferroni correction. In the second motion experiment there were no significant differences in RT or accuracy across experimental conditions (*t*-test, $P > 0.2$), except that accuracy was higher for motion-holes than for motion-surfaces, Bonferroni corrected (Table 2). Overall,

TABLE 2. Behavioral responses during scan

Experiment 1	Object	Hole	Surfaces	
(a) Stereo				
Number correct, %	92 ± 10	90 ± 11	90 ± 13	
RT, ms	753 ± 73	752 ± 82	761 ± 80	
(b) Motion				
Number correct, %	96 ± 5	96 ± 6	98 ± 4	
RT, ms	733 ± 113	782 ± 127	785 ± 141	
Experiment 2	Object	Hole	Edges	Surfaces
(a) Stereo				
Number correct, %	96 ± 3	92 ± 5	93 ± 5	94 ± 4
RT, ms	797 ± 51	771 ± 102	754 ± 62	749 ± 85
(b) Motion				
Number correct, %	90 ± 6	95 ± 3*	89 ± 6	83 ± 6
RT, ms	807 ± 141	808 ± 127	784 ± 95	788 ± 134

Values are means ± SD. An asterisk indicates significantly different from other conditions ($P < 0.05$). Accuracy during *experiment 2* was higher for motion-holes versus other conditions; however, only the difference between motion-holes and motion-surfaces remained significant after Bonferroni correction.

performance was similar across conditions and there was no condition in which performance was consistently better or worse.

LOC responds preferentially to objects across multiple visual cues

The lateral occipital complex (LOC; Grill-Spector et al. 2001; Malach et al. 1995) consists of a constellation of regions that respond more strongly to objects than to scrambled objects or random stimuli. We used an independent localizer experiment to find in each subject regions that activate more strongly to gray-level images of objects than to scrambled versions of these objects (METHODS). Object-selective activations occurred in the lateral bank of the occipital and temporal cortex, extending anteriorly into the fusiform gyrus. In all subjects, regions that responded preferentially to gray-level images of objects (animals, cars, and novel) than to scrambled objects ($P < 0.005$, voxel level, Fig. 3A) also responded more strongly to stereo-objects than to random-disparity stereograms ($P < 0.005$, voxel level, Fig. 3B) and more strongly to motion-objects than to random motion ($P < 0.005$, voxel level, Fig. 3C). These data concur with prior reports of object-selective responses in the LOC across multiple visual cues (Gilaie-Dotan et al. 2002; Grill-Spector et al. 1998; Kastner et al. 2000; Kourtzi and Kanwisher 2000; Mendola et al. 1999).

On average, a larger extent of lateral occipital cortex was activated by stereo-objects than to motion-objects [right hemisphere: stereo, $3,812 \pm 2,378 \text{ mm}^3$ (mean ± SD); motion, $1,957 \pm 1,266 \text{ mm}^3$; left hemisphere: stereo, $3,959 \pm 2,418 \text{ mm}^3$; motion, $1,688 \pm 1,333 \text{ mm}^3$]. We found substantial overlap in the spatial extent of regions activated more strongly to stereo-objects than to random-stereo and more strongly to motion-objects than to random motion. The extent of overlap was limited by the smaller size of the region that showed higher responses for motion objects [% overlap, right hemisphere: (motion-objects and stereo-objects)/motion-objects = $70 \pm 27\%$, (motion-objects and stereo-objects)/stereo-objects = $39 \pm 23\%$; left hemisphere: (motion-objects and stereo-objects)/motion-objects = $64 \pm 31\%$, (motion-objects and stereo-objects)/stereo-objects = $37 \pm 25\%$].

LOC processes shape, not edges or surfaces

We extracted responses from LOC subregions defined by the gray-level localizer experiment (Figs. 2 and 3 and METHODS) during motion and stereo experiments to examine whether the LOC is involved in processing of shape, local edges, or global surfaces across multiple cues. Using a four-way ANOVA we examined the effects of *stimulus type* (*experiment 1*: object/hole/surfaces/random; *experiment 2*: object/hole/edges/surfaces/random), *cue* (stereo/motion), *ROI* [LO (a), (b); pFUS/OTS], and *hemisphere* (left/right) on fMRI responses. Each cell in the ANOVA contained the mean response for a condition (across six blocks) for each subject, and subjects were treated as a random variable.

We found a significant effect of stimulus type [*experiment 1*: $F_{(3,519)} = 48.06$, $P < 10^{-6}$; *experiment 2*: $F_{(4,389)} = 18.24$, $P < 10^{-6}$], indicating that LOC responses differ across object, hole, edges, surfaces, and random. There was no significant effect of cue in *experiment 2* [$F_{(1,389)} = 0.57$, $P = 0.45$] and no significant effect of cue when the ANOVA was performed

across both experiments [$F_{(1,509)} = 1.01, P = 0.39$], but there was a small effect of cue in *experiment 1* [$F_{(1,519)} = 7.54, P = 0.01$]. Moreover, there were no significant interactions between cue and other factors in either *experiment 1* or *experiment 2* ($F < 2.24; P > 0.07$), suggesting that LOC responses are similar across motion and stereo stimuli. We found a main effect of ROI [*experiment 1*: $F_{(4,519)} = 40.84, P < 10^{-6}$; *experiment 2*: $F_{(2,389)} = 20.37; P < 10^{-6}$]. Responses were not different between hemispheres [*experiment 1*: $F_{(4,519)} = 4.44, P = 0.0356$; *experiment 2*: $F_{(2,389)} = 0.86; P = 0.35$, except for LO (b) in *experiment 1*] and there were no significant interactions between any of the factors and hemisphere. Therefore subsequent analyses were performed separately for each ROI, averaged across hemispheres.

We next examined whether processing of either surfaces or edges alone contribute to LOC responses. Across all LOC ROIs we found no significant difference between the responses to surfaces and responses to the random condition in either motion or stereo experiments [paired *t*-test, $t(12) < 0.88, P > 0.2$; Fig. 4]. This suggests that global surfaces without edges (generated from motion or disparity) do not differentially activate LOC compared with random stimuli. For both motion and stereo cues disconnected edges did not activate LOC ROIs significantly more than surfaces [Fig. 4B, $t(7) < 1.7, P > 0.06$] or more than random [Fig. 4B, $t(7) < 0.58, P > 0.29$].

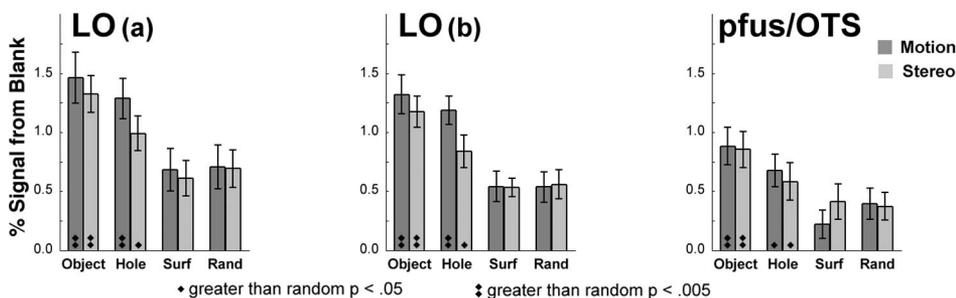
However, for both motion and stereo cues, LOC subregions responded to objects and holes significantly more than to random stimuli [Fig. 4C, objects: $t(12) > 5.14, P < 0.0002$; holes: $t(12) > 2.1, P < 0.03$] and significantly more than to global surfaces [Fig. 4C, objects: $t(12) > 4.2, P < 10^{-3}$; holes: $t(12) > 2.6, P < 0.02$, except for pFUS/OTS, stereo holes vs. stereo surfaces, $t(12) = 1.6, P = 0.07$]. Activation to objects was also significantly higher than that to edges [Fig. 4C, $t(7) > 3.16, P < 0.01$] in both motion and stereo experiments.

These data indicate that the LOC responds more strongly to objects than to global surfaces, disconnected edges, or random stimuli for both motion and stereo cues. In contrast, global segmentable surfaces without edges do not activate LOC more than random stimuli. Finally, edges that did not enclose a global shape did not activate LOC regions more than random stimuli or global surfaces.

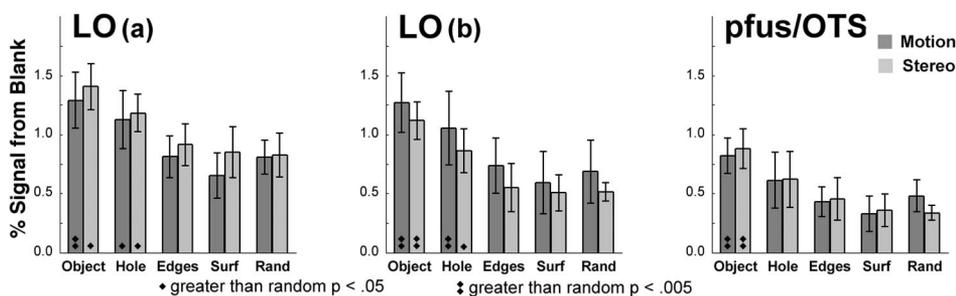
LOC responds more strongly to objects than to holes

Notably, LOC's response to objects was greater than that to holes for both motion and stereo cues [Fig. 4C, paired *t*-test, $t(12) > 2.73; P < 0.01$], even though the contour that defined the shape of objects and holes was identical. Since the only difference between objects and holes is the border ownership of the contour defining the shaped region, this suggests that LOC is sensitive to border ownership. Alternatively, higher

A Experiment 1



B Experiment 2



C Objects vs. other conditions

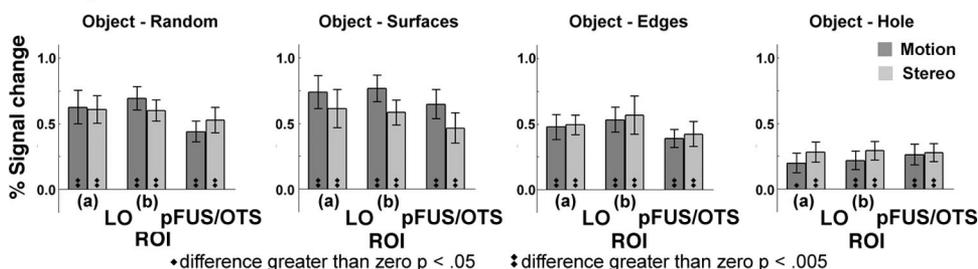


FIG. 4. LOC responses during stereo and motion experiments. *A*: average blood oxygen level-dependent (BOLD) response in LOC ROIs (defined from the gray-level localizer experiment) during motion *experiment 1* (dark gray) and stereo *experiment 1* (light gray) across 9 subjects. % Signal is plotted relative to the blank (fixation) baseline. *B*: average responses from LOC ROIs (defined from the gray-level localizer experiment) in motion and stereo *experiment 2* averaged across 8 subjects. In *A* and *B*, diamonds reflect significantly higher responses in a condition than the random condition (paired *t*-test across subjects). *C*: differential LOC responses between objects and other conditions, averaged across all 13 subjects (except for object-edges that include 8 subjects from *experiment 2*). Diamonds reflect differences that are significantly greater than zero (paired *t*-test across subjects). Error bars indicate SE across subjects. Surf, surfaces; Rand, random.

LOC responses for objects than for holes may be a consequence of lower responses to holes than to objects in early visual areas or may relate to differences in perceiving a shaped region during hole versus object conditions. We examined each of these possibilities in turn.

Processing of surfaces, edges, and shape in early and intermediate visual areas

We extracted responses from retinotopic visual regions (METHODS and Fig. 2) to examine whether differential responses in LOC reflect downstream responses propagated from early visual cortex and to test whether retinotopic regions show cue-specific or cue-independent preferential responses to edges or global surfaces compared with random stimuli. Single-unit recordings suggest that luminance-edges and motion-edges are processed in early visual areas (Hubel 1995) and that responses of V1, V2, and V4 neurons are modulated by the border ownership of edges (Qiu and von der Heydt 2005; Zhou et al. 2000).

V1–V3 activated robustly to all experimental conditions including the random condition (Fig. 5, A–C). There were no significant differences in V1 responses across stimulus types [object/hole/edges/surfaces/random, $F_{(4,69)} = 0.86, P = 0.49$] or across motion and stereo stimuli [$F_{(1,69)} = 3.56, P = 0.06$, same for *experiment 1*; data not shown]. V2 and V3 responses were not significantly different across experimental conditions [$F_{(4,69)} < 0.48; P > 0.75$], although responses to motion stimuli were higher than those to stereo stimuli [$F_{(1,69)} > 6.3, P < 0.02$]. The lack of differential response across experimental conditions in early visual cortex validates that our experimental conditions were controlled for low-level visual properties. There were no significant differences in responses of early retinotopic regions between objects and holes, for either motion or stereo cues [$t(10) < 1.7; P > 0.06$]. This suggests that higher responses in LOC to objects than to holes were not propagated from early visual cortex.

Intermediate visual areas V3a and V4 showed no significant effects of cue or stimulus type ($F < 2.79, P > 0.1$). Notably, V3a and V4 activations to stimuli containing motion-edges (edges, hole, and object conditions) were significantly higher than those to random motion (each comparison, paired *t*-test, $P < 0.03$; Fig. 5, D and E) and motion-surfaces (Fig. 5, D and E; paired *t*-test, $P < 0.03$). Responses were not different across motion-objects, motion-holes, or motion-edges ($P > 0.35$). Activation to stereo-edges was not significantly higher than that to random-disparity stereograms or stereo-surfaces ($P > 0.38$).

MT responses were significantly higher for motion than for stereo stimuli [$F_{(1,79)} = 39.51; P < 10^{-6}$]. Although the main effect of stimulus type was not significant [$F_{(4,79)} = 1.21; P = 0.32$], MT responded significantly more strongly to motion-objects and motion-holes than to random motion or motion-surfaces, and more to stereo-objects than to random stereo or stereo-surfaces [$t(11) > 2.34, P < 0.02$; Fig. 5F]. Thus MT showed a strong preference for motion stimuli, with a slight preference for objects (see also Kourtzi et al. 2002).

Overall, analyses of retinotopic visual areas found no difference in their responses for objects versus holes, indicating that the preference to objects observed in LOC was not propagated from lower visual areas. We found evidence for the involvement of V3a and V4 in motion-edge processing, evidence for general processing of visual motion in MT, and a weak preference for motion in V2 and V3.

Subjects perceive a shaped region in both object and hole conditions

We next examined whether subjects were able to perceive a shaped region during hole and object conditions (see METHODS) to determine whether differential LOC responses for objects and holes related to differential perception of the shaped region. Subjects participated in a behavioral experiment in which they viewed the same objects and holes as those during fMRI scans, generated from either motion or stereo cues, and

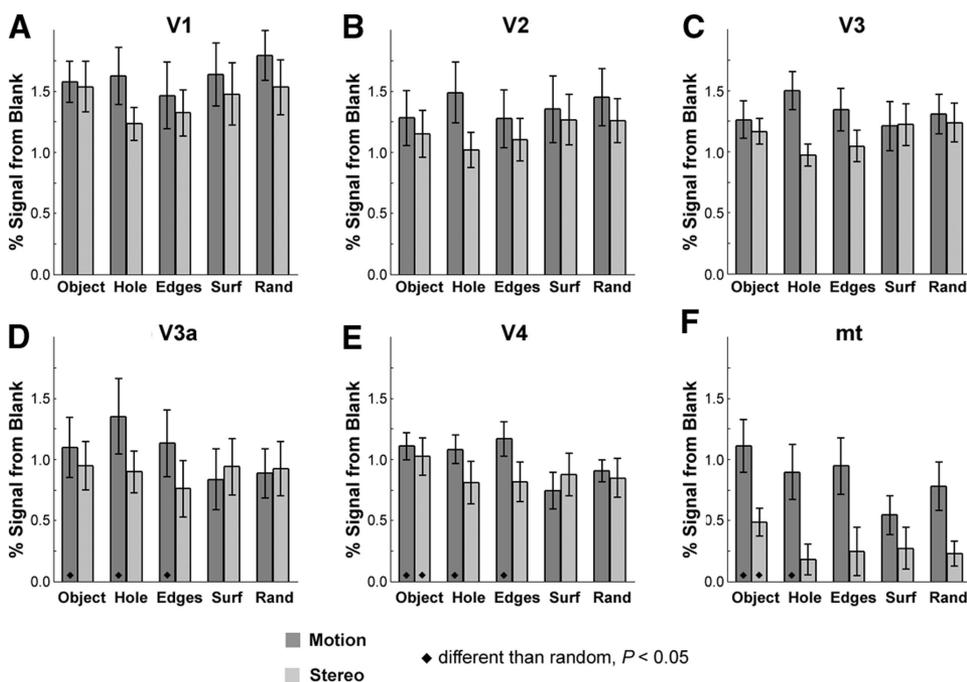


FIG. 5. Activation of retinotopic visual areas and MT. Mean BOLD response amplitudes relative to a blank (fixation) baseline during motion and stereo *experiment 2* for retinotopic visual areas (7-subject average) and MT (8-subject average). Error bars indicate SE across subjects.

were instructed to indicate whether the fixation point was inside or outside the shape. This task examines whether subjects can determine a shaped region in the image (Stanley and Rubin 2003). Lower accuracy on this task for holes would indicate that subjects were worse at determining the shaped region for holes versus objects.

Subjects' accuracy was high for both objects and holes, indicating that subjects were able to perceive a shaped region for both object and holes (percentage accuracy for motion-objects, $97 \pm 3\%$; motion-holes, $98 \pm 3\%$; stereo-objects, $97 \pm 2\%$; stereo-holes, $93 \pm 8\%$). There were no significant differences in accuracy (paired *t*-test, all values of $P > 0.07$, uncorrected). Subjects responded more quickly to objects than to holes (motion-objects, 967 ± 108 ms; motion-holes, $1,025 \pm 115$ ms; stereo-objects, 902 ± 128 ms; stereo-holes, $1,056 \pm 240$ ms, $P < 0.005$). The longer RTs for holes are unlikely to explain the lower LOC responses for holes since, during scanning, there were no RT differences (Table 2) and conditions with longer response times are usually associated with higher LOC signals (Sayres and Grill-Spector 2006), not lower responses. Since subjects were able to perceive the shaped region for both objects and holes, the lower LOC activation for holes is unlikely to be due to subjects' inability to determine the shaped region for hole stimuli.

Voxel-by-voxel analyses of selective responses to shape, edges, and global surfaces

To complement ROI analyses we performed a voxel-by-voxel search for any voxels that activated more strongly to objects, holes, edges, or surfaces than to random (Fig. 6). These analyses were done on a subject-by-subject basis. The goal of these analyses was to detect any region in visual cortex that showed preferential responses to edges, global surfaces, or global shapes than to random stimuli containing no structure for each of the visual cues.

Voxel-based analyses revealed that higher activation to shape versus random (object vs. random and hole vs. random, $P < 0.01$, voxel level) in both motion and stereo experiments occurred in higher-order visual regions including LOC, but excluded early visual cortex (Fig. 6, *A* and *B*). These activations overlapped with LOC ROIs in all subjects. Ventral LOC activation to holes versus random was less robust than objects versus random for both motion and stereo cues in all subjects: Fewer voxels in anterior ventral regions were significantly activated by the former versus latter contrast (Fig. 6*B*). Higher activation for holes than that for random ($P < 0.01$, voxel level) overlapped most subjects' LO [ROI (a): stereo: LH, 12/13; RH, 10/13; motion: LH, 12/13; RH 9/13; ROI (b): stereo: LH, 12/13; RH 12/13; motion: LH, 9/13; RH, 12/13], but in fewer subjects in pFUS/OTS (stereo: LH, 7/13; RH, 7/13; motion: LH, 3/13; RH, 3/13).

Voxel-by-voxel activations for edges versus random ($P < 0.01$, voxel level, Fig. 6*C*) revealed a dorsal region adjacent to V3a on the lateral bank of the intraparietal sulcus and/or middle occipital gyrus that activated more strongly both to motion-edges versus random motion and also to stereo-edges versus random disparity (Fig. 6*C*, Talairach coordinates in Table 1*B*). This pattern of response occurred bilaterally in six of eight subjects and provides evidence for cue-invariant processing of edges. Examination of retinotopic maps indicates

that these preferential responses to edges for both motion and stereo stimuli overlapped the retinotopic representation of V7. We will therefore refer to this ROI as V7. Another activation focus that responded more strongly to edges than to random was detected in the cIPS: cIPS responded more strongly for stereo-edges than for random stereo in all subjects ($P < 0.01$, voxel level, RH: 8/8 subjects; LH: 6/8 subjects) and in half the subjects also to motion-edges versus random motion ($P < 0.01$, voxel level, RH: 5/8 subjects; LH: 4/8 subjects).

We analyzed the mean response across V7 (Fig. 7*A*; see METHODS). ANOVA of V7 responses with factors of stimulus type (object/hole/edges/surfaces/random), cue (motion/stereo), and hemisphere (right/left) revealed a significant effect of stimulus type [$F_{(4,119)} = 7.1$, $P < 0.00005$], but no effect of cue or interactions ($F < 1.27$, $P > 0.26$). Responses in the right hemisphere were stronger than those in the left [$F_{(1,119)} = 9.23$, $P < 0.003$]. Since there was no significant interaction between hemisphere and condition, responses were averaged across hemispheres. V7 responded significantly more strongly to edges versus random, holes versus random, and objects versus random for both motion and stereo stimuli (paired *t*-test, $P < 0.01$; Fig. 7*A*). In contrast, responses to either motion-surfaces or stereo-surfaces were not higher than those to random ($P > 0.49$). Responses to random disconnected edges were not different from responses to holes or objects that contained a shape and responses to objects were not different from holes ($P > 0.22$). This suggests that V7 participates in processing of edges from both motion and stereo cues, irrespective of whether they define a global shape.

We also found evidence for cue-specific processing of edges: higher activation to motion-edges versus random motion occurred in V3a (5/7 subjects), V4 (6/7 subjects), and LO (a, 5/8 subjects), consistent with ROI analyses (Fig. 5).

Cue-specific processing of surfaces in IPS regions

Voxel-by-voxel analyses of global surfaces versus random stimuli ($P < 0.01$, voxel level) showed that voxels in the intraparietal sulcus activated more strongly to stereo-surfaces than to random-disparity stereograms. Activations occurred in two foci: in cIPS (RH: 9/13, LH: 8/13 subjects) and in an IPS region between V7 and cIPS (RH: 7/13, LH: 6/13 subjects; Fig. 6*D*, *left*). In contrast, few voxels responded significantly more strongly to motion-surfaces than to random motion (Fig. 6*D*, *right*). Even when we reduced our statistical threshold to $P < 0.05$ (voxel level, uncorrected) we found little activations for motion-surfaces versus random motion. When they occurred, they partially overlapped MT (5/13 subjects), dorsal V3 (2/11 subjects), or V3a (4/11 subjects), rather than IPS regions. However, these differential responses were not significant across subjects (all values of $t < 0.2$; $P > 0.45$; Fig. 5). Thus we did not find evidence for any visual region that shows preferential responses to global surfaces relative to the random condition across motion and stereo cues.

We extracted responses from IPS regions (see METHODS). Responses in the IPS region between V7 and cIPS (Fig. 7*B*) showed no main effect of cue, stimulus type, or hemisphere ($F < 2.03$, $P > 0.09$) or any interactions ($F < 0.45$, $P > 0.77$). Responses were higher for stereo-surfaces than for random stereo [$t(6) > 2.1$, $P < 0.04$], but were not higher for

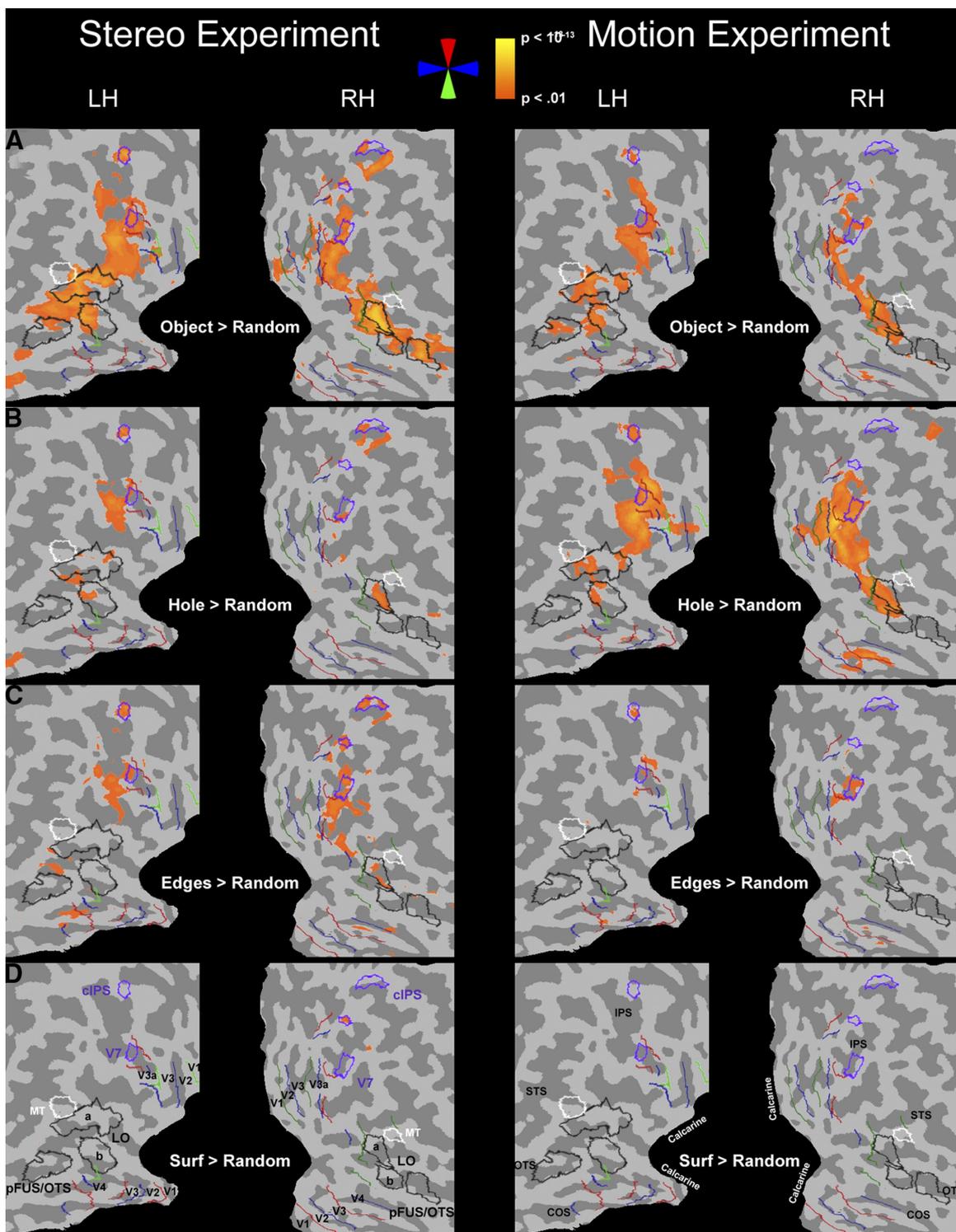


FIG. 6. Voxel-by-voxel statistical maps of activations compared with random. Voxel-by-voxel activation maps of each condition vs. random in each of the stereo and motion experiments were generated for each subject and displayed on flattened hemispheres, with a representative subject shown here. Statistical threshold is identical across all maps. Black, LOC ROIs: LO (a), (b), pFUS/OTS from the gray-level localizer experiment; white, MT; purple, dorsal ROIs along the IPS, from posterior to anterior; V7, DF2: a region between V7 and cIPS; and cIPS. Visual meridians: red, upper; blue, horizontal; and green, lower. Caudal intraparietal sulcus (cIPS); COS, collateral sulcus; IPS, intraparietal sulcus; OTS, occipitotemporal sulcus; pFUS, posterior fusiform gyrus; STS, superior temporal sulcus.

motion-surfaces than for random motion [$t(6) = -0.69$, $P = 0.74$].

cIPS responses (Fig. 7C) revealed a significant effect of stimulus type [$F_{(4,139)} = 3.93$, $P < 0.005$], cue [$F_{(1,139)} = 4.9$,

$P < 0.03$], and hemisphere [$F_{(1,139)} = 4.8$, $P < 0.03$], but there were no significant interactions ($P > 0.23$). Responses were higher for stereo than for motion stimuli and higher in the right than in the left hemisphere. During the stereo experiment

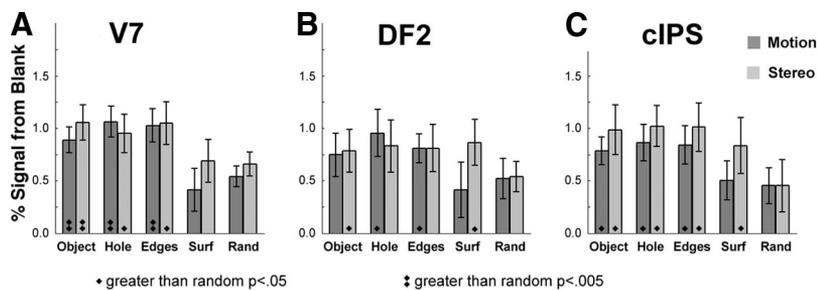


FIG. 7. Mean BOLD responses in dorsal ROIs. Mean BOLD responses are plotted relative to the blank (fixation) baseline for motion and stereo *experiment 2*. *A*: average V7 responses across 6 subjects. *B*: average responses for an IPS region between V7 and cIPS across 6 subjects. *C*: average cIPS responses across 8 subjects. Error bars indicate SE across subjects.

responses were significantly higher for surfaces versus random, edges versus random, holes versus random, and objects versus random-disparity stereograms [$t(7) > 2.9$, $P < 0.01$; Fig. 7C]. However, responses for motion-surfaces were not significantly higher than those for random motion [$t(7) = 0.6$, $P = 0.37$]. There were no significant differences across stereo-objects, holes, edges, or surfaces [$t(7) < 1.9$, $P > 0.06$]. Thus the higher responses for global surfaces compared with random occurred only for stereo stimuli. Overall, cIPS responded more strongly to stereo than to motion stimuli and both stereo-surfaces and stereo-edges activated this region more than did random-disparity stereograms.

DISCUSSION

Shape selectivity across multiple cues in the LOC

Our data provide strong evidence for selectivity for shape-from-luminance, shape-from-motion, and shape-from-stereo in the LOC: the LOC responded more strongly to object and holes than to random-dot stimuli, disconnected edges, and global surfaces without edges across motion and stereo cues. However, LOC's responses to edges or global surfaces were not different from random stimuli across both stereo and motion cues. These data provide strong evidence that LOC does not represent global surfaces or local edges. LOC's similar activation for shapes defined from either motion or stereo cues may reflect either invariance at the single-neuron level or invariance across the population of LOC neurons. That is, cue invariance revealed by standard fMRI may reflect a homogeneous cue-invariant neural population or a mixture of heterogeneous neural populations selective to shape from different cues within fMRI voxels. Findings of cue-invariant responses to shapes in macaque inferior temporal neurons (Sary et al. 1993; Tanaka et al. 2001; Vogels and Orban 1996) and results of fMRI-adaptation studies in humans (Kourtzi and Kanwisher 2000) favor the former interpretation. Future fMRI-adaptation studies in humans are necessary to determine whether the higher responses to motion objects and stereo objects versus control stimuli involve the same neural population or distinct neural populations in the LOC. Overall, our data underscore the involvement of LOC in processing shape across multiple visual cues.

Processing of edges across multiple cues in the human visual cortex

We found evidence for both cue-specific and cue-independent processing of edges. Selective responses to edges across motion and stereo stimuli overlapped the visual field map named V7. V7 responded more strongly to edges than to

random stimuli and surfaces for both motion and stereo cues. Activation to edges, objects, and holes was similar, suggesting that disconnected edges, which do not form a global shape, are sufficient to robustly activate this region. This extends previous findings (Gilaie-Dotan et al. 2002; Grill-Spector et al. 1998) that implicated this dorsal region in processing objects by suggesting that local edges alone are sufficient to activate it more than random stimuli. These selective responses to edges across multiple cues may be useful for detecting contours in the visual input.

Cue-specific processing of edges occurred in V3a and V4: these regions responded more strongly to motion-edges than either to random motion or to motion-surfaces, but they did not respond more strongly to stereo-edges than to random stereo. Further, these regions responded more strongly to any stimulus that contained a motion edge (objects or hole) than to random stimuli. Since the higher responses for edges occurred only for one cue, it suggests a cue-specific response to motion edges.

Cue-specific representation of global surfaces

We did not find any region that showed selective responses to global surfaces versus random across motion and stereo cues. However, we found evidence for stereo-specific processing of global surfaces along the IPS. Notably, cIPS responded more strongly to stereo-surfaces than to random-disparity stereograms and also more strongly to stereo than to motion stimuli. These results are consistent with previous studies suggesting the role of cIPS in processing stereo cues in humans and monkeys (Shikata et al. 2001; Taira et al. 2000; Tsao et al. 2003; Tsutsui et al. 2005). Our data suggest that cIPS is involved in processing stereo information and it also prefers stereo-surfaces and stereo-edges over random-disparity stereograms.

It is possible that cIPS responses to stereo-surfaces were higher than motion-surfaces because of the greater perceived depth during the stereo experiments. However, this explanation is unlikely because random-disparity stereograms and stereo-surfaces had similar depth range, yet responses to stereo-surfaces were higher than responses to random stereograms. Another possibility is that small differences in eye movements across conditions may have contributed to differential responses in parietal activations (Schluppeck et al. 2005). However, it is not obvious why differential eye movements would occur only for stereo experiments. Future experiments examining responses to surfaces defined from other cues that create a powerful depth perception such as texture gradients (Tsutsui et al. 2002) and measurements of eye movements will determine whether surface processing in cIPS is specific to stereo information.

A question that remains is why did we not in general find evidence for robust responses to global surfaces than to random stimuli across visual cortex? One possibility is that regions involved in surface representation are involved in both local and global surface computations and therefore activate similarly for global surfaces and the random conditions (that may have local surface information). Another possibility is that the mean response across a visual region may not be higher for global surfaces than for random stimuli or may not differ across various kinds of surfaces, but within these regions there are neurons that are sensitive to surface information. fMRI adaptation, for example, may reveal sensitivity to surface orientation or depth even when the mean response to surfaces of different orientations or depths is similar. Future experiments are required to distinguish between these possibilities and for a more comprehensive understanding of surface processing in visual cortex.

Comparison of fMRI data to single-unit studies of early visual cortex and MT

Analyses of low-level visual areas revealed that there were no significant differences between conditions, validating that experimental conditions were matched for low-level properties. Our finding of shape-selective activations in LOC across multiple cues, but not early visual cortex, concurs with several neuroimaging studies (Appelbaum et al. 2006; Gilaie-Dotan et al. 2002; Grill-Spector et al. 1998; Schira et al. 2004) and an electrophysiology study (Rossi et al. 2001). However, these results differ from electrophysiology studies that reported that V1 and V2 neurons respond more vigorously when a cell's receptive field overlaps an object boundary or lies within a shaped-figure region (Lamme 1995; Qiu and von der Heydt 2005; Zhou et al. 2000; Zipser et al. 1996). One possibility for the discrepancy between neuroimaging and electrophysiology data is that modulations of neural responses during figure-ground segmentation are small (relative to the large initial nonspecific response) and occur in a minority of neurons. Therefore these modulations cannot be observed by fMRI that measures the pooled activity across a neural population. Alternatively, our random stimuli had more local variance in the range of disparity or motion than the object or hole conditions and therefore produced more activation in V1 neurons that have small receptive fields. A third explanation, based on recurrent models of figure-ground segmentation (Jehee et al. 2007; Roelfsema et al. 2002), suggests that processing of the figure region occurs across multiple visual areas. According to this model, the area that will show the largest modulations during figure-ground computations is the region in which neurons' receptive fields match the size of the figure region. Since we used shapes that were about 11° , the largest effects would be observed in higher-order areas (such as LOC) that contain neurons with relatively large receptive fields (Yoshor et al. 2007). Supporting this idea, electrophysiology data show that figure-ground effects in V1 occur only when the figure region is small ($<4^\circ$) and its edge is close to the cell's receptive field (Rossi et al. 2001; Zipser et al. 1996).

We also found evidence for preferential responses to motion in MT, V2, and V3. These regions responded in general more strongly to motion stimuli than to stereo stimuli. Of the regions we measured, MT showed the largest cue-specific responses.

To what extent can we attribute higher MT responses to moving than to stationary stereo stimuli to a cue effect? Although our stimuli were matched for dot density, contrast, contour length and performance was similar across stereo and motion experiments (Table 2), it is not possible to equate X units of motion to Y units of stereo. Nevertheless, higher MT responses for moving stimuli than for stationary stereo may be explained by the properties of MT neurons. In monkeys 96% of MT neurons are sensitive to the direction of motion, 99% to the speed of motion, and 93% to horizontal disparity (DeAngelis and Uka 2003; DeAngelis et al. 1998; Palanca and DeAngelis 2003). Most experiments that investigated MT responses to stereo used moving stimuli, except for the Palanca and DeAngelis (2003) study. When Palanca and colleagues examined neural responses to moving stereo stimuli versus stationary stereo stimuli they found that 66% of MT neurons showed a significant reduction in their response to stationary stereo stimuli compared with moving stereo stimuli (to a level less than one third of the response to moving stimuli). Since fMRI signals likely reflect the overall neural responses across space and time, lower firing rates for stationary than for moving stimuli are likely to underlie the lower MT responses to stationary stereo stimuli. Nevertheless, a future quantitative study relating neural firing and fMRI responses for moving and stationary stereo stimuli is necessary to validate this explanation.

Role of LOC in shape computations

We found higher LOC responses for objects than for random, surfaces, and edges across motion and stereo cues. Importantly, we found no difference in LOC responses to global surfaces compared with random stimuli containing no structure for either motion or stereo cues. In a previous study, Stanley and Rubin (2003) created modified Kanizsa shapes via "pacmen-like" inducers in which there was a clear percept of an enclosed region even when the bounding illusory contours were not perceived. Because they found that the LOC responded more strongly to enclosed regions than to similar stimuli (created by rotating the pacmen), which did not enclose a region, they concluded that the presence of a global surface is sufficient to activate LOC even when its illusory edges were not perceived. However, our data show that global surfaces that are not enclosed do not activate LOC more than random stimuli, and that global shapes activate LOC more than global surfaces. Therefore our data suggest that the pattern of LOC responses reported by Stanley and colleagues may have been driven by the processing of the global shape of the enclosed region rather than the processing of its global surface.

Our findings of higher LOC responses for stereo-objects versus stereo-edges and higher responses for motion-objects versus motion-edges (created by randomly scrambling object contours) extend previous results that reported higher LOC responses to luminance-defined global shapes versus luminance-defined random contours, misaligned contours, or partial object contours (Altmann et al. 2003; Doniger et al. 2000; Lerner et al. 2002, 2004). Although the amount of dots that moved (or were in the near disparity) in the edges condition was lower than that in the object, hole, or surfaces conditions, it is unlikely that this factor explains lower LOC responses to edges. There were more dots that moved (or had near disparity) in the random or surfaces conditions than in the object or hole

conditions, but this did not produce higher LOC responses. Overall, our data provide evidence that LOC also processes global shape information (rather than local edges) for motion and stereo cues.

Notably, we found that LOC activated more strongly to objects than to holes across both motion and stereo cues. This higher response to objects than to holes occurred specifically in the LOC and more prominently in ventral subregions, which responded robustly to objects versus random, but did not respond robustly to holes versus random (Fig. 6B). We did not find differential responses between holes and objects in other visual regions. What is the source of this differential response? One possibility is that this pattern of responses reflects a bias for near disparity. However, this explanation is unlikely because both object and hole conditions had the same relative disparity and the contour defining the shape was in the nearer surface in both conditions. Alternatively, lower LOC responses to holes may reflect worse recognition of the shape of holes than the shape of objects because reduced LOC responses are correlated with lower accuracy during object recognition tasks (Bar et al. 2001; Grill-Spector et al. 2000; James et al. 2000). Although we did not find performance differences between objects and holes in our experiments, differences may become apparent during tasks that require finer discriminations (e.g., distinguishing between a cow and a deer). A third possibility is that the difference between responses to objects and holes reflects the difference in the assignment of the border ownership of the contour defining the shape. In the object condition, the border is an intrinsic boundary and “belongs to the object,” whereas the border defining the hole is an extrinsic contour that “belongs” to the front occluding surface. This explanation suggests that information about border ownership may be processed in the LOC. Nevertheless, explanations based on border ownership and recognition are not mutually exclusive. It is possible that the LOC participates in border ownership computations and is optimized to process shapes defined by intrinsic contours (objects) rather than shapes defined by extrinsic contours (holes). Therefore LOC responses are higher for objects than for holes and, as a consequence, the perception of objects may be more salient than the perception of holes. Further research is necessary to distinguish between these alternatives.

Implications for theories of figure–ground segmentation

Our results have important implications for theories of figure–background segmentation. Figure–ground segmentation is the process in which the figure region (a shaped front region) is segmented from the rest of the image (the shapeless “background”). Traditional theories of figure–ground segmentation propose a bottom-up computation, in which determining surfaces and their relative depth occurs before shape processing and this process also determines the border ownership of edges (Nakayama 1995; Nakayama et al. 1989). Nakayama and colleagues suggested that representation of global surfaces occurs in a visual region between V1 and inferotemporal cortex. Other theories suggest that figure–ground segmentation includes a top-down computation in which shape information is used during figure–ground processing (Peterson and Gibson 1993, 1994; Roelfsema et al. 2002; Vecera and O’Reilly 1998).

Several aspects of our data challenge traditional bottom-up theories of figure–ground segmentation. First, bottom-up theories suggest an intermediate region in which global surfaces are represented. However, we did not find evidence for cue-invariant representation of global surfaces. Second, for both motion and stereo cues, LOC activation to objects was higher than that to holes, but there was no difference in activation to objects and holes in other visual regions. This suggests that computations of border ownership (or computations of the figure region) may occur in higher stages of the visual hierarchy, such as regions involved in object and shape perception, contrary to predictions of bottom-up theories. Therefore our data suggest the involvement of object-selective cortex during figure–ground segmentation (Peterson and Gibson 1993, 1994; Roelfsema et al. 2002; Vecera and O’Reilly 1998).

In sum, we found evidence for selective responses to global shape across multiple visual cues in the LOC, evidence for both cue-specific and cue-independent responses to edges in intermediate visual areas, and stereo-specific responses to global surfaces in IPS regions. Further, our results suggest that some types of information, specifically motion-edges and stereo-surfaces, may be represented in a cue-specific manner in distinct cortical regions. Future experiments testing additional visual cues (e.g., color or texture) will advance our understanding of the nature of cue-specific processing in these regions. Overall, these data indicate that integration across multiple visual cues is mainly achieved at the level of shape and underscore LOC’s role in shape computations.

ACKNOWLEDGMENTS

We thank D. Andersen, B. Dougherty, D. Remus, R. Sayres, N. Witthoft, and J. Winawer for fruitful discussions and for comments on earlier versions of this manuscript.

GRANTS

This research was funded by Whitehall Foundation Grant 2005-05-111-RES to K. Grill-Spector and a Stanford Undergraduate Research Opportunities grant to J. Vinberg.

REFERENCES

- Albright TD, Desimone R, Gross CG.** Columnar organization of directionally selective cells in visual area MT of the macaque. *J Neurophysiol* 51: 16–31, 1984.
- Altmann CF, Bulthoff HH, Kourtzi Z.** Perceptual organization of local elements into global shapes in the human visual cortex. *Curr Biol* 13: 342–349, 2003.
- Appelbaum LG, Wade AR, Vildavski VY, Pettet MW, Norcia AM.** Cue-invariant networks for figure and background processing in human visual cortex. *J Neurosci* 26: 11695–11708, 2006.
- Bar M, Tootell RB, Schacter DL, Greve DN, Fischl B, Mendola JD, Rosen BR, Dale AM.** Cortical mechanisms specific to explicit visual object recognition. *Neuron* 29: 529–535, 2001.
- Castelo-Branco M, Formisano E, Backes W, Zanella F, Neuenschwander S, Singer W, Goebel R.** Activity patterns in human motion-sensitive areas depend on the interpretation of global motion. *Proc Natl Acad Sci USA* 99: 13914–13919, 2002.
- DeAngelis GC, Cumming BG, Newsome WT.** Cortical area MT and the perception of stereoscopic depth. *Nature* 394: 677–680, 1998.
- DeAngelis GC, Uka T.** Coding of horizontal disparity and velocity by MT neurons in the alert macaque. *J Neurophysiol* 89: 1094–1111, 2003.
- Doniger GM, Foxe JJ, Murray MM, Higgins BA, Snodgrass JG, Schroeder CE, Javitt DC.** Activation timecourse of ventral visual stream object-recognition areas: high density electrical mapping of perceptual closure processes. *J Cogn Neurosci* 12: 615–621, 2000.
- Dupont P, De Bruyn B, Vandenberghe R, Rosier AM, Michiels J, Marchal G, Mortelmans L, Orban GA.** The kinetic occipital region in human visual cortex. *Cereb Cortex* 7: 283–292, 1997.

- Ffytche DH, Skidmore BD, Zeki S.** Motion-from-hue activates area V5 of human visual cortex. *Proc Biol Sci* 260: 353–358, 1995.
- Gilade-Dotan S, Ullman S, Kushnir T, Malach R.** Shape-selective stereo processing in human object-related visual areas. *Hum Brain Mapp* 15: 67–79, 2002.
- Glover GH.** Simple analytic spiral K-space algorithm. *Magn Reson Med* 42: 412–415, 1999.
- Goodale MA, Milner AD.** Separate visual pathways for perception and action. *Trends Neurosci* 15: 20–25, 1992.
- Grill-Spector K, Kourtzi Z, Kanwisher N.** The lateral occipital complex and its role in object recognition. *Vision Res* 41: 1409–1422, 2001.
- Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzhak Y, Malach R.** Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron* 24: 187–203, 1999.
- Grill-Spector K, Kushnir T, Edelman S, Itzhak Y, Malach R.** Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron* 21: 191–202, 1998.
- Grill-Spector K, Kushnir T, Hendler T, Malach R.** The dynamics of object-selective activation correlate with recognition performance in humans. *Nat Neurosci* 3: 837–843, 2000.
- Grill-Spector K, Malach R.** The human visual cortex. *Annu Rev Neurosci* 27: 649–677, 2004.
- Heeger DJ, Boynton GM, Demb JB, Seidemann E, Newsome WT.** Motion opponency in visual cortex. *J Neurosci* 19: 7162–7174, 1999.
- Hubel DH.** *Eye, Brain and Vision*. New York: Freeman, 1995.
- Huk AC, Heeger DJ.** Pattern-motion responses in human visual cortex. *Nat Neurosci* 5: 72–75, 2002.
- James TW, Humphrey GK, Gati JS, Menon RS, Goodale MA.** The effects of visual object priming on brain activation before and after recognition. *Curr Biol* 10: 1017–1024, 2000.
- Jehee JF, Lamme VA, Roelfsema PR.** Boundary assignment in a recurrent network architecture. *Vision Res* 47: 1153–1165, 2007.
- Kastner S, De Weerd P, Ungerleider LG.** Texture segregation in the human visual cortex: a functional MRI study. *J Neurophysiol* 83: 2453–2457, 2000.
- Kourtzi Z, Bulthoff HH, Erb M, Grodd W.** Object-selective responses in the human motion area MT/MST. *Nat Neurosci* 5: 17–18, 2002.
- Kourtzi Z, Kanwisher N.** Cortical regions involved in perceiving object shape. *J Neurosci* 20: 3310–3318, 2000.
- Lamme VA.** The neurophysiology of figure–ground segregation in primary visual cortex. *J Neurosci* 15: 1605–1615, 1995.
- Larsson J, Heeger DJ.** Two retinotopic visual areas in human lateral occipital cortex. *J Neurosci* 26: 13128–13142, 2006.
- Lerner Y, Harel M, Malach R.** Rapid completion effects in human high-order visual areas. *Neuroimage* 21: 516–526, 2004.
- Lerner Y, Hendler T, Malach R.** Object-completion effects in the human lateral occipital complex. *Cereb Cortex* 12: 163–177, 2002.
- Malach R, Reppas JB, Benson RR, Kwong KK, Jiang H, Kennedy WA, Ledden PJ, Brady TJ, Rosen BR, Tootell RB.** Object-related activity revealed by functional magnetic resonance imaging in human occipital cortex. *Proc Natl Acad Sci USA* 92: 8135–8139, 1995.
- Marr D.** *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*. New York: Freeman, 1982.
- Mendola JD, Dale AM, Fischl B, Liu AK, Tootell RB.** The representation of illusory and real contours in human cortical visual areas revealed by functional magnetic resonance imaging. *J Neurosci* 19: 8560–8572, 1999.
- Mishkin M, Ungerleider LG, Macko KA.** Object vision and spatial vision: two cortical pathways. *Trends Neurosci* 6: 414–417, 1983.
- Nakayama K, He ZJ, Shimojo S.** Visual surface representation: a critical link between lower-level and higher-level vision. In: *Visual Cognition* (2nd ed.), edited by Osherson DN, Kosslyn SM. Cambridge, MA: MIT Press, 1995.
- Nakayama K, Shimojo S, Silverman GH.** Stereoscopic depth: its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception* 18: 55–68, 1989.
- Newsome WT, Britten KH, Movshon JA.** Neuronal correlates of a perceptual decision. *Nature* 341: 52–54, 1989.
- Palanca BJ, DeAngelis GC.** Macaque middle temporal neurons signal depth in the absence of motion. *J Neurosci* 23: 7647–7658, 2003.
- Peterson MA, Gibson BS.** Shape recognition contributions to figure–ground organization in three-dimensional displays. *Cogn Psychol* 25: 383–429, 1993.
- Peterson MA, Gibson BS.** Must shape recognition follow figure–ground organization? *Psychol Sci* 5: 253–259, 1994.
- Press WA, Brewer AA, Dougherty RF, Wade AR, Wandell BA.** Visual areas and spatial summation in human visual cortex. *Vision Res* 41: 1321–1332, 2001.
- Qiu FT, von der Heydt R.** Figure and ground in the visual cortex: V2 combines stereoscopic cues with Gestalt rules. *Neuron* 47: 155–166, 2005.
- Roelfsema PR, Lamme VA, Spekreijse H, Bosch H.** Figure–ground segregation in a recurrent network architecture. *J Cogn Neurosci* 14: 525–537, 2002.
- Rossi AF, Desimone R, Ungerleider LG.** Contextual modulation in primary visual cortex of macaques. *J Neurosci* 21: 1698–1709, 2001.
- Sack AT, Kohler A, Linden DE, Goebel R, Muckli L.** The temporal characteristics of motion processing in hMT/V5+: combining fMRI and neuronavigated TMS. *Neuroimage* 29: 1326–1335, 2006.
- Sakata H, Taira M, Kusunoki M, Murata A, Tsutsui K, Tanaka Y, Shein WN, Miyashita Y.** Neural representation of three-dimensional features of manipulation objects with stereopsis. *Exp Brain Res* 128: 160–169, 1999.
- Sary G, Vogels R, Orban GA.** Cue-invariant shape selectivity of macaque inferior temporal neurons. *Science* 260: 995–997, 1993.
- Sayres R, Grill-Spector K.** Object-selective cortex exhibits performance-independent repetition suppression. *J Neurophysiol* 95: 995–1007, 2006.
- Schira MM, Fahlke M, Donner TH, Kraft A, Brandt SA.** Differential contribution of early visual areas to the perceptual process of contour processing. *J Neurophysiol* 91: 1716–1721, 2004.
- Schluppeck D, Glimcher P, Heeger DJ.** Topographic organization for delayed saccades in human posterior parietal cortex. *J Neurophysiol* 94: 1372–1384, 2005.
- Seidemann E, Poirson AB, Wandell BA, Newsome WT.** Color signals in area MT of the macaque monkey. *Neuron* 24: 911–917, 1999.
- Shikata E, Hamzei F, Glauche V, Knab R, Dettmers C, Weiller C, Buchel C.** Surface orientation discrimination activates caudal and anterior intraparietal sulcus in humans: an event-related fMRI study. *J Neurophysiol* 85: 1309–1314, 2001.
- Stanley DA, Rubin N.** fMRI activation in response to illusory contours and salient regions in the human lateral occipital complex. *Neuron* 37: 323–331, 2003.
- Taira M, Tsutsui KI, Jiang M, Yara K, Sakata H.** Parietal neurons represent surface orientation from the gradient of binocular disparity. *J Neurophysiol* 83: 3140–3146, 2000.
- Tanaka H, Uka T, Yoshiyama K, Kato M, Fujita I.** Processing of shape defined by disparity in monkey inferior temporal cortex. *J Neurophysiol* 85: 735–744, 2001.
- Tootell RB, Reppas JB, Kwong KK, Malach R, Born RT, Brady TJ, Rosen BR, Belliveau JW.** Functional analysis of human MT and related visual cortical areas using magnetic resonance imaging. *J Neurosci* 15: 3215–3230, 1995.
- Tsao DY, Vanduffel W, Sasaki Y, Fize D, Knutsen TA, Mandeville JB, Wald LL, Dale AM, Rosen BR, Van Essen DC, Livingstone MS, Orban GA, Tootell RB.** Stereopsis activates V3A and caudal intraparietal areas in macaques and humans. *Neuron* 39: 555–568, 2003.
- Tsutsui K, Sakata H, Naganuma T, Taira M.** Neural correlates for perception of 3D surface orientation from texture gradient. *Science* 298: 409–412, 2002.
- Tsutsui K, Taira M, Sakata H.** Neural mechanisms of three-dimensional vision. *Neurosci Res* 51: 221–229, 2005.
- Tyler CW, Likova LT, Kontsevich LL, Wade AR.** The specificity of cortical region KO to depth structure. *Neuroimage* 30: 228–238, 2006.
- Van Oostende S, Smaert S, Van Hecke P, Marchal G, Orban GA.** The kinetic occipital (KO) region in man: an fMRI study. *Cereb Cortex* 7: 690–701, 1997.
- Vecera SP, O'Reilly RC.** Figure–ground organization and object recognition processes: an interactive account. *J Exp Psychol Hum Percept Perform* 24: 441–462, 1998.
- Vogels R, Orban GA.** Coding of stimulus invariances by inferior temporal neurons. *Prog Brain Res* 112: 195–211, 1996.
- Yoshor D, Bosking WH, Ghose GM, Maunsell JH.** Receptive fields in human visual cortex mapped with surface electrodes. *Cereb Cortex* 17: 2293–2302, 2007.
- Zeki S.** Thirty years of a very special visual area, Area V5. *J Physiol* 557: 1–2, 2004.
- Zhou H, Friedman HS, von der Heydt R.** Coding of border ownership in monkey visual cortex. *J Neurosci* 20: 6594–6611, 2000.
- Zipser K, Lamme VA, Schiller PH.** Contextual modulation in primary visual cortex. *J Neurosci* 16: 7376–7389, 1996.