Beyond Retinotopic Mapping: The Spatial Representation of Objects in the Human Lateral Occipital Complex

Ayelet McKyton and Ehud Zohary

1Neurobiology Department, Life Science Institute and
2Interdisciplinary Center for Neural Computation,
Hebrew University, Jerusalem 91904, Israel

The spatial representation in the human ventral object-related areas (i.e., the lateral occipital complex [LOC]) is currently unknown. It seems plausible, however, that it would diverge from the strict retinotopic mapping (characteristic of V1) to a more invariant coordinate frame, thereby allowing for reliable object recognition in the face of eye, head, or body movement. To study this, we compared the fMRI activation in LOC when object displacement was limited to either the retina or the screen by manipulating eye position and object locations. We found clear adaptation in LOC when the object's screen position was fixed, regardless of the object's retinal position. Furthermore, we found significantly greater activation in LOC in the hemisphere contralateral to the object's screen position, although the visual task was constructed in a way that the objects were present equally often on each of the 2 retinal hemifields. Together, these results indicate that a sizeable fraction of the neurons in LOC may have head-based receptive fields. Such an extraretinal representation may be useful for maintenance of object coherence across saccadic eye movements, which are an integral part of natural vision.

Keywords: coordinate transformation, egocentric, fMRI, object recognition, vision

Introduction

The spatial representation of information in the earliest phases of visual processing (retina, Lateral Geniculate Nuclei [LGN]) is in a retinotopic coordinate frame. Obviously, this is insufficient for veridical perception: Eye movements cause displacement of a static object’s image on the retina, although it remains stationary in the world. On the other hand, tracking a moving object with the eyes, as done in smooth pursuit, results in displacement of the object in the world without concomitant movement on the retina. Resolution of this problem can be obtained by incorporation of information about the eye position.

Von Helmholtz (1860, reproduced in 1962) suggested that a copy of the motor command, that is, the corollary discharge, might be sent to the visual areas to allow disambiguation of retinal motion from real movement in the world. In monkeys, it has been known for some time that the eye’s position in the orbit can modulate the activity of a sizeable fraction of neurons in the parietal cortex (Andersen and Mountcastle 1983; Andersen and others 1985, 1990; Andersen 1997) as well as in occipital areas such as V3A (Galletti and Battaglini 1989; Nakamura and Colby 2000). One common view is that this reflects a gradual coordinate transformation toward an egocentric representation that takes place in the dorsal pathway, in accordance with its putative role in guiding action (Andersen and Buneo 2002). Indeed, in area Ventral Intraparietal Area (VIP) of the parietal cortex, there is a functional heterogeneity among neurons: the receptive fields of some neurons map the location of the stimulus in space, irrespective of the monkey’s direction of gaze (i.e., head-centered mapping), whereas others demonstrate the conventional retinotopic (eye position dependent) receptive fields (Duhamel and others 1997). Other parietal neurons are affected by both factors (showing eye position-dependent gain fields). Recent modeling efforts showed that such neuronal characteristics (gain fields, head-centered receptive fields) can emerge naturally from an attractor network connectivity and dynamics (Pouget and others 2002).

Evidence of eye position-dependent modulation of neuronal activity has also been reported in ventral areas in the monkey (V4/V8: Dobbins and others 1998; Bremner 2000; inferotemporal cortex [IT]: Nowicka and Ringo 2000) and even in some neurons in V1 (Trotter and others 1992, 1996; Li and Guo 1997; Trotter and Celebrini 1999). Because a saccade is usually toward the most salient visual feature in the scene, eye movements are inherently related to the focus of attention (though the 2 can obviously be uncoupled). Spatial attention is clearly important for object recognition and modulates the activity of neurons in the ventral stream (Moran and Desimone 1985). Thus, it may not be too surprising that microstimulation of neuronal populations involved in planning eye movements in the frontal eye fields at an amplitude that does not generate a saccade, can still affect the neuronal firing in V4 (Moore and Armstrong 2003), or that eye movements can modulate visual receptive fields of V4 neurons (Tolias and others 2001).

Thus, it is plausible that object-related areas in the ventral stream may incorporate an eye position signal to generate a more invariant representation than a retinotopic one. One possibility is an extraretinal (head centered) representation in which the neurons’ firing rate is maintained in spite of changes in gaze direction if the stimulus is kept in the same position relative to the head. Such an extraretinal representation seems especially useful to maintain a coherent image of an object (or the visual scene) in the presence of often occurring saccadic eye movements, which are an integral part of natural vision.

We therefore designed an experiment aimed at studying the spatial coordinate frame in ventral object-related areas: the lateral occipital complex (LOC). LOC is situated lateral and anterior to V4/V8 and includes the lateral occipital sulcus, the fusiform gyrus, and the collateral sulcus (Malach and others 1995, 2002; Grill-Spector and others 1999). Because the spatial resolution of functional magnetic resonance imaging (fMRI) is limited, such that the signal reflects the average activity of large groups of neurons (~10⁶), we applied the functional magnetic resonance adaptation (fMR-A) technique (Grill-Spector and Malach 2001) to indirectly examine the functional properties of groups of neurons in LOC. One of the key characteristics of...
LOC is that the hemodynamic response to a sequence of images containing the same repeated stimulus is smaller than to a sequence of different images (i.e., LOC shows stimulus-specific adaptation). Grill-Spector and Malach reasoned that if changing a visual property in the repeated stimulus causes recovery of the hemodynamic response (i.e., recovery from adaptation), then the neurons in this region, at least at the population level, are sensitive to this property. This has some corroborative evidence from single-unit studies in anterior IT in monkeys, which show such adaptation patterns (Li and others 1993; Lueschow and others 1994). Thus, by testing if adaptation can be found when the stimulus is stationary on the screen, but changes position on the retina (due to eye movements), we could probe the functional properties of groups of neurons in LOC.

However, it is unclear whether LOC is the human homologue of anterior IT, and the interpretation of fMR-A results in neuronal terms is indirect and somewhat controversial. It is therefore crucial to confirm the interpretation based on the adaptation results using a different approach. We introduce a novel approach based on the fact that LOC shows a clear contralateral preference, displaying greater activation when the stimulus is shown on the contralateral visual field than when it is presented in the ipsilateral one (Niemeier and others 2005). We utilize this contralateral preference by creating a situation in which the stimuli are shown in opposite visual hemifields for the 2 coordinate frames. For example, a stimulus is presented so that it will be on the “right” part of the “screen” (i.e., in the right hemifield in head-centered coordinates) but on both sides of the “retina” (due to eye movements that ensure this). Comparison of the fMRI signal in the 2 hemispheres, individually for each subject, could provide evidence whether neurons in the region typically encode information in a retinal- or head-based coordinate frame, especially when analyzed in tandem with the adaptation results.

Materials and Methods

Magnetic Resonance Imaging Subjects
A total of 10 volunteers without neurological, psychiatric, or visual deficits history (6 women and 4 men aged 25–40 years) participated in the present experiments. The Tel-Aviv Sourasky Medical Center Ethic Committee approved the experimental procedure. A written informed consent was obtained from each subject.

Magnetic Resonance Imaging Acquisition
The blood oxygenation level-dependent fMRI measurements were performed in a whole-body 1.5-T, Sigma Horizon, LX8.25 General Electric scanner. The fMRI protocols were based on a multislice gradient echo-planar imaging and a standard head or surface coil. The functional data were obtained under the optimal timing parameters: time repetition = 3 s, time echo = 55 ms, flip angle = 90°, imaging matrix = 128 × 128, and field of view = 24 cm. The 27 slices with slice thickness 4 mm (with no gap) were oriented in the axial position. In-plane resolution was 3 × 3 mm. The scan covered the whole brain. The fMRI images were superimposed on T1-weighted 3-dimensional (3D) SPGR images (3D-fast spoiled GRASS, spatial resolution: 1 × 1 × 1 mm).

Experimental Setup
The visual sequences were generated on a PC and projected via LCD projector (Epson MP-7200, Tokyo, Japan) onto a tangent screen located inside the scanner in front of the subject. Subjects viewed the screen through a tilted mirror.

Stimuli and Experimental Paradigms

LOC Localizer
A set of 64 black and white pictures of objects and 64 scrambled pictures of the same objects were used. A fixation point (0.3 degrees) appeared in the middle of the screen (20.3 degrees) throughout the experiment. The objects were man-made tools/devices, and their size was roughly 3 degrees. The experiment was carried out using a block design format. There were 4 block types (see Fig. 1, top): Objects centered 4 degrees left (and 0.8 degrees above) of the fixation point, objects 4 degrees right (and 0.8 degrees below) of the fixation point, and scrambled objects at the same locations. Each block lasted 12 s followed by a blank period of 9 s. The first and last blank periods were longer (27° and 15 s, respectively). There were 24 objects or scrambled objects in each block, each presented for 500 ms. Subjects were instructed to maintain fixation throughout the experiment and during epochs to covertly name the object. A short training procedure was applied before the experiment to assure that the subjects recognize all the objects.

Experiment 1
Experiment 1 included the same pictures of man-made tools/devices presented in the same screen locations as in the localizer experiment (see Fig. 3a). In all conditions, the subjects performed a covert recognition task while fixating on the center of the screen. All epochs lasted 12 s followed by 9 s of rest period. Twelve objects were presented in each block, each for 800 ms followed by 200 ms in which only the fixation point was present. However, unlike the localizer, the pictures were repeated within a block (using either 4 pictures, repeated 3 times [Fig. 3a, fixed4 and alter4 conditions], or 2 pictures, repeated 6 times [fixed2 and alter2 conditions]) to get differential adaptation effects. Furthermore, in 2 conditions (fixed2 and fixed4), the pictures were always shown on the right part of the screen, whereas in the other 2 conditions (alter4 and alter2), the objects changed their position, alternating between the left and right side of the screen.

Experiment 2
Experiment 2 included the same pictures of man-made tools/devices, presented in the same screen locations as in experiment 1 (see Fig. 4a). The block length and timing of each stimulus were the same as in experiment 1. The pictures were also repeated within a block (using either 4 pictures, repeated 3 times [Fig. 4a, fixed-head4 and fixed-retina4 conditions], or 2 pictures, repeated 6 times [fixed-head2 and fixed-retina2 conditions]), or 2 pictures, repeated 6 times [fixed-head2 and fixed-retina4 conditions]). Similarly, in 2 conditions (fixed-head2 and fixed-head4), the pictures were always shown on the right part of the screen.
Eye movements were recorded outside the magnet in 4 subjects while they were performing eye movements as in experiment 2. The measurements were taken using an infrared eye tracker (EyeLinkI) in the same context as during the scan (identical visual stimuli and task as in experiment 2). Generally, the subjects were able to keep fixation well within a window of 1 degree. The standard deviation of the jitter was 0.95 degrees across subjects and an average of 0.39 within subjects, much smaller than the objects' positions located approximately 4 degrees left or right of the fixation point. No significant difference in eye position was observed between the “fixed retina” and the “fixed-head” conditions (see Supplementary Fig. 1).

Eye Movement Measurement

Eye movements were recorded outside the magnet in 4 subjects while they were performing eye movements as in experiment 2. The measurements were taken using an infrared eye tracker (EyeLinkI) in the same context as during the scan (identical visual stimuli and task as in experiment 2). Generally, the subjects were able to keep fixation well within a window of 1 degree. The standard deviation of the jitter was 0.95 degrees across subjects and an average of 0.39 within subjects, much smaller than the objects' positions located approximately 4 degrees left or right of the fixation point. No significant difference in eye position was observed between the “fixed retina” and the “fixed-head” conditions (see Supplementary Fig. 1).

Data Analysis

Data analysis was performed using the BrainVoyager 4.96 and BrainVoyager QX software package (Brain Innovation, Maastricht, The Netherlands, 2000). For each subject, the 2D functional data were aligned to 2D anatomical slices of the same subject. Before statistical analysis, raw data were examined for motion and signal artifacts. Head motion correction and high-pass temporal filtering in the frequency domain were applied in order to remove drifts and to improve the signal-to-noise ratio. Time courses were obtained from the LOC region of interest (ROI). Voxels were selected to within the LOC ROI if they showed a significant activation for objects (left and right to the fixation point) compared with scrambled objects (left and right to the fixation point) using the general linear model analysis. Only voxels within the occipitotemporal cortex with q(FDR) smaller than 0.01 were chosen (Genovese and others 2002). This procedure ensures that the chance for a “false” inclusion of a voxel in the ROI is no greater than 1%. The obtained maps were superimposed on the same individual’s 3D anatomical reference scan. The 3D recordings were used for surface reconstruction. This procedure included the segmentation of the white matter using a growing region function. The cortical surface was then unfolded, cut along the calcaneous sulcus, and flattened. The obtained activation maps were superimposed on an inflated and unfolded cortical map for each subject (as in Fig. 1). Time courses were taken from the LOC ROI. The subject’s average activation was estimated by averaging the activation during the period between 6 and 12 s after epoch onset across all the voxels in the LOC ROI. Finally, average signal intensity across all subjects was calculated, pooling over the different individual average signal intensities.

Results

We investigated the patterns of cortical activation in 10 subjects using fMRI during 2 block-designed experiments. Each experiment consisted of 4 conditions repeated 5 times in a counterbalanced fashion. A rest condition, lasting 9 s, that followed each epoch served as a hemodynamic baseline condition. Each epoch, lasting 12 s, consisted of either 2 or 4 different pictures of parfoveally displayed objects, which were repeatedly shown in a cyclic manner (Figs 3a and 4a). The subject’s head was directed toward the center of the screen.

We began by using an external localizer paradigm to delineate LOC as the ROI. The localizer included pictures of objects and scrambled versions of the same objects presented left or right of a fixation point using a block design paradigm (see Materials and Methods). Voxels located in the occipitotemporal cortex, which showed significantly greater activation for the original objects compared with the scrambled ones (either in the left or in the right side), were selected individually for each subject (Fig. 1). The localizer experiment was necessary to verify the existence of 2 characteristics of LOC, which were a prerequisite for the interpretation of our results in the following experiments.

The first feature, “contralateral bias,” was demonstrated in the past in LOC (Niemeyer and others 2005). These results were repeated using our paradigm: activation in the hemisphere contralateral to the stimulus was greater than in the hemisphere ipsilateral to the stimulus (Fig. 2b). Activation in the right hemisphere (RH) was significantly higher than in the left hemisphere (LH) during presentation of the objects on the “left” side ($t$-test: $P < 7 \times 10^{-7}$). The opposite picture was observed when the objects were presented on the right side ($t$-test: $P < 4 \times 10^{-7}$).

The second tested feature was “hemispheric symmetry”: Symmetry in the pattern of activation in the 2 hemispheres when comparing the activation evoked by presenting a contralateral stimulus to each hemisphere (i.e., RH activation during presentation of the objects on the left side vs. LH activation during presentation of the objects on the right side). Indeed, the activation in LOC was very similar in the 2 hemispheres (Fig. 2c): RH and LH activations were similar during presentation of a contralateral or an ipsilateral stimulus. There was no significant difference between RH and LH activations to a contralateral object ($t$-test: $P = 0.42$) or an ipsilateral stimulus ($t$-test: $P = 0.09$).

Furthermore, the average number of activated voxels was very similar in both hemispheres (LH 7870 voxels; RH 7481 voxels; $P = 0.48$).

To summarize, LOC has a contralateral bias showing a much greater activation for a stimulus presented in the contralateral side than in the ipsilateral side. LOC also seems to show hemispheric symmetry having similar fMRI activation in the 2 hemispheres for contralateral (or ipsilateral) stimuli.

Experiment 1

The localizer experiment showed that altering the position of the stimulus by 8 degrees across the midline markedly changes...
the pattern of activation in LOC. Experiment 1 was designed to verify that the activation in LOC is specific to the object's position as indicated from the localizer experiment. The experiment used 2 different methods: Adaptation and hemisphere comparison. These methods will be essential to study the spatial coordinate frame in LOC (see Experiment 2). Subjects were scanned maintaining fixation at a stationary fixation point, while pictures of objects appeared in the periphery in 2 different ways (see Fig. 3a). Either objects appeared always to the right of the fixation point in a fixed location ("fixed" conditions) or the objects appeared to the left and right of the fixation point in an alternating manner ("alter" conditions).

Figure 3. Experimental design and results of experiment 1. (a) The stimuli used were black and white pictures of objects presented at 4 degrees to the left or to the right of a stationary fixation point (marked by cross). The numbers on the left upper corner were not shown on the screen and represent here the specific object identity. Subjects were instructed to maintain fixation and covertly name each object. All epochs lasted 12 s and contained 12 images of objects each presented for 800 ms with a 200-ms interstimulus interval. Four experimental conditions were used: fixed4—4 objects repeating 3 times were shown to the right of the fixation point, fixed2—2 objects repeating 6 times in the same position, alter4—4 objects were alternated between the right and left screen positions, alter2—2 repeated objects were shown, alternating in position as in alter4. (b) Hemodynamic signal in the LH and RH for the 4 different conditions in LOC, averaged across 10 subjects. Gray background represents the epoch period (12 s). (c) The average activation, across all subjects, showing adaptation between the fixed conditions and not between the alter conditions. (d) The average activation, across all subjects, showing contralateral bias: higher activation in the contralateral LH during the fixed conditions and "symmetry": no difference in activation between the hemispheres during the alter conditions. Error bars represent standard error of mean.
Adaptation

During the fixed and the alter conditions, either 4 or 2 repeated objects were presented in different epochs (see Fig. 3a). The degree of fMRI adaptation during presentation of 2 objects compared with the presentation of 4 objects was assessed both in the fixed conditions (fixed4 and fixed2) and in the alter conditions (alter4 and alter2). Based on findings from previous studies (Grill-Spector and Malach 2001), we expected to find adaptation in the fixed2 condition compared with fixed4. This is due to the fact that the fixed2 condition was constructed from less different objects than fixed4. Furthermore, in these conditions, the repeated objects appeared in a fixed location. A similar adaptation profile in the alter conditions, in spite of the altering position of the objects, would suggest that the receptive fields of neurons in LOC are large enough, penetrating ~4 degrees into the ipsilateral side. On the other hand, a recovery from adaptation in the alter conditions would suggest that different populations of neurons in LOC are active when the stimulus is shown in the 2 different locations.

Indeed, adaptation was observed only between the fixed conditions, whereas a total recovery from adaptation was present between the alter conditions (see Fig. 3c). Thus, fMRI activation during fixed2 was significantly smaller than during fixed4 conditions in the LH (contralateral) (t-test, LH: P < 0.005). No such adaptation was evident in the comparison between the alternating conditions alter2 and alter4 (alter2 activation was not significantly different than alter4, t-test, LH: P = 0.98, RH: P = 0.56).

Hemisphere Comparison

Next, we applied a method, termed hemisphere comparison, to corroborate our findings from the adaptation analysis using a different methodology. To that end, we compared the fMRI activation during the alter conditions, in which objects were presented in both hemisfields, and during the fixed conditions, in which objects were presented only in the right hemifield (Fig. 3a). We reasoned that a stronger hemodynamic signal in the LH in the fixed conditions (in which the objects are on the right side) will indicate, once again, the existence of contralateral bias in the representation in LOC. This was indeed the case (Fig. 3d, contralateral bias) (RH activation was significantly smaller than LH activation during the fixed conditions, t-test, fixed4: P < 10−5, fixed2: P < 10−5). Furthermore, similar activation between the hemispheres was found in the alter conditions, confirming the symmetry between hemispheres seen in the localizer experiment (Fig. 3d, symmetry; t-test, alter4: P = 0.48, alter2: P = 0.08).

In conclusion, the results from both adaptation and hemisphere comparison analysis indicate that neurons in LOC are specific to the stimulus’ location, showing a major change in their response properties when the visual image is displaced across the vertical meridian by ~8 degrees (~4 degrees deep into each hemifield). The results (from both this experiment and the localizer experiment) indicate that the receptive fields of neurons in LOC typically represent the contralateral visual field (thereby generating the contralateral bias in the fMRI signal) and that the representation of objects is symmetric between the 2 hemispheres.

There are at least 2 alternative explanations as to why recovery from adaptation would be evident when the same object is presented in different positions in space. The classic explanation is that cells in the LOC have a restricted “retinotopic” receptive field, such that the visual stimulus excites nonoverlapping neuronal populations as it is shown in different screen locations (corresponding to different retinal locations). However, due to the fact that the eyes remained stationary throughout the experiment, a specific position on the screen corresponded to a given location on the retina. Thus, an alternative hypothesis that the neurons in LOC have restricted “world,” “body,” or “head-based” receptive fields is also consistent with these results. According to this explanation, the receptive field of neurons in LOC is defined in head-based coordinates (or body or even world-based coordinates) so that the neurons respond to a specific place with respect to the head (rather than to a specific position on the retina). To determine if the response of neurons in LOC is according to a retinotopic-based coordinate frame or to head (body or world)-based coordinates, it is crucial to generate a condition in which changes are induced in one coordinate framework without concomitant change in the other. This was studied in experiment 2.

Experiment 2

Experiment 2 was designed to reveal whether the specificity to location that was found in experiment 1 is in a retinotopic or in a head-based coordinate system. To do so, we created conditions in which only the retinal position of the object changed with no concurrent changes in the object’s position on the screen. This was achieved by alternating the fixation point between 2 screen positions while the object’s position on the screen remained stationary. We termed these conditions the fixed-head conditions (Fig. 4a, fixed-head conditions). In contrast, in the fixed-retina conditions, the object’s position on the screen changed with minimal changes on the retina due to eye movements that maintained the same geometric relations between the fixation point and object’s position (Fig. 4a, fixed-retina conditions).

Adaptation

During the fixed-head and the fixed-retina conditions, either 4 or 2 repeated objects were presented in different epochs (see Fig. 4a). The degree of fMRI adaptation during presentation of 2 objects compared with the presentation of 4 objects was assessed both in the fixed-head conditions (fixed head4 and fixed head2) and in the fixed-retina conditions (fixed retina4 and fixed retina2).

We reasoned that if the representation in LOC is retinotopic, one should observe adaptation when comparing the fixed retina2 with the fixed retina4 because both activated the same retinal location. Retinotopic representation would further predict no adaptation when comparing fixed head2 with fixed head4 because the repeated objects in these conditions are shown in “opposite” retinal hemifields (see Fig. 4a; note that the results from experiment 1 indicate that typically the receptive fields do not stretch out to include both the contralateral and ipsilateral parafocal regions). However, if the receptive fields in LOC are head based (or body or even world based), the opposite picture is expected: Adaptation should be observed between the 2 fixed-head conditions (because the repeated objects’ position is stationary on the screen and relative to the head) and not between the fixed-retina conditions.

As predicted by a head-based representation, adaptation was observed only between the fixed-head conditions (see Fig. 4c; fixed-head2 activation was significantly smaller than fixed-head4, t-test, LH: P < 0.017, RH: P < 0.010). On the other
hand, near complete recovery from adaptation was present between the fixed-retina conditions (fixed-retina2 activation was not significantly different than fixed-retina4, \( t \)-test, LH: \( P = 0.78 \), RH: \( P = 0.59 \)).

**Hemisphere Comparison**

Next, we compared the fMRI activation in the fixed-retina conditions and the fixed-head ones in the 2 hemispheres. We reasoned that if the representation in LOC was a retinotopic...
one, a higher hemodynamic response should be seen in the contralateral RH during the fixed-retina conditions (compared with the ipsilateral LH). This is because in the fixed-retina conditions, the objects were always present in the left retinal hemifield (see Fig. 4a). On the other hand, in the fixed-head conditions, the activation in the 2 hemispheres is expected to be similar because during these epochs the objects were equally present in the 2 retinal hemifields.

A completely different picture is expected if the representation is in head-based coordinates. In this case, no difference between hemispheres should be evident in the fixed-retina conditions because the objects appear in both sides of the head (i.e., with respect to the head direction). In contrast, during the fixed-head conditions, a higher fMRI signal would be expected in the LH because the objects appeared always to the right of the head.

The results show both effects: During the fixed-retina conditions, in which the objects always appeared to the left of the fixation point (but on both sides of the head), we found a greater fMRI signal in the RH than in the left one (see Fig. 4d, contralateral bias to retinal position; t-test, fixed retina: \( t < 0.007 \), fixed retina: \( P < 0.001 \)). This result suggests that some neurons in LOC maintain their representation in retinotopic coordinates.

However, during the fixed-head conditions, in which objects appeared to the “right of the head” but in both retinal hemifields, a greater fMRI signal was evident in the LH than in the right one (see Fig. 4d, contralateral bias to head position; t-test, fixed head: \( P < 10^{-5} \), fixed head: \( P < 0.001 \)). This effect was even greater than the contralateral preference seen in the fixed-retina conditions. This result suggests the existence of neurons in LOC whose representation is in head-based coordinates.

To conclude, the results from both the adaptation and hemisphere comparison analyses indicate that the representation of a sizeable proportion of neurons in LOC is different from the classical retinotopic one in that it is dependent on gaze angle.

**Discussion**

The aim of this research was to study the nature of the spatial representation of objects in LOC. Using the adaptation technique and our novel analysis method, utilizing the contralateral preference of the visual areas, our experiments reveal 2 important properties about the ventral object-related areas. First, the fMRI signal in LOC is sensitive to object location suggesting that the effective receptive fields of most neurons in the human object-related areas do not extend into the ipsilateral side of the visual field in more than 4 degrees (experiment 1). Second, and most important, the fMRI signal is sensitive to the object’s position on the screen even more than to the object’s position on the retina (experiment 2). This suggests that the receptive fields of a sizeable fraction of neurons in LOC are not organized according to a retinotopic principle. Rather, they seem to encode object location in at least head-centered coordinates. This profile of activation was found throughout LOC (including the collateral sulcus, the fusiform, or the lateral occipital sulcus; data not shown). This suggests that a representation in an extraretinal coordinate frame is common to LOC as a whole.

**Possible Confounding Effects: Eye Movements and Screen Edges**

Could the current results be due to possible confounding factors such as inaccurate eye movements or the presence of screen edges? It is important to realize that the required eye movements were identical in all conditions in experiment 2. Thus, to get a differential pattern of activation in the different conditions, one would have to posit that the subjects were making consistently different eye movements during the different conditions. This seems highly unlikely. Still, we tested the accuracy of saccadic eye movements while the subjects were performing a task, as in experiment 2, outside the scanner. Our results indicate that the eye movements were quite accurate and similar across conditions (see Supplementary data).

One might also posit that in experiment 2, the screen edges could cause a recovery from adaptation because the right edge of the screen appeared closer to the fovea every second picture. However, because this “edge effect” was common to “all conditions,” it cannot explain any difference between conditions and certainly not the lack of adaptation in the retina conditions. We therefore conclude that our main finding that the fMRI signal in LOC is sensitive to the object’s position on the screen, more than to the object’s position on the retina, is unlikely to be due to such artifacts.

**A Mosaic of Spatial Frameworks in LOC**

Our finding that the fMRI signal in LOC is largely dependent on the position of the objects on the screen does not imply that none of the neurons in LOC are retinotopically tuned. A more plausible explanation of our data is that there are heterogeneous populations of neurons in LOC data (similar to findings in area VIP of the monkey; Duhamel and others 1997). One is strictly retinotopic and responsible for the retinotopic contralateral bias and maybe to some adaptation (though not significant) between the fixed-retina conditions (Fig. 4). The other population is extraretinotopic and responsible for the head-based contralateral bias and the adaptation seen during the fixed-head conditions. It is important to note that the contralateral bias to the head position is more robust than the contralateral bias to retinal position (Fig. 4d). This could suggest a larger extraretinotopic population than a retinotopic one. This, together with the finding that adaptation effects are smaller than the contralateral bias effects, can explain why the adaptation between the fixed-retina conditions was not significant.

**LOC and IT Receptive Fields**

How does our finding fit in with the current literature on the receptive field structure in the monkey? Typically, the inferior temporal cortex (IT) is considered as the putative analogue of LOC in the monkey, serving as the ventral object recognition complex (Gross and others 1979; Mishkin and Ungerleider 1982). A recent extensive investigation of neurons in area TE (which is located within IT) shows that they have a mean receptive field size of 9.6 degrees (range: 2.8–25.9 degrees, for half the maximum response) with a strong preference for the contralateral field (Op De Beeck and Vogels 2000). Specifically, about 63% of the neurons had their receptive field centered on the fovea, 34% on the contralateral field, and only 1.4% on the ipsilateral field, outside the fovea. Furthermore, a recent study (DiCarlo and Maunsell 2003) showed that neurons in anterior IT have fine position sensitivity, showing a dramatic reduction in their firing rates for object displacements as small as 1.5 degrees, especially when the object is placed in the ipsilateral visual field. These findings fit well the results in experiment 1, in which the fMRI signal in the ipsilateral (right) hemisphere was roughly 3 times smaller than the signal in the contralateral one.
A recent fMRI adaptation experiment in humans (Grill-Spector and others 1999), focusing on the ventral object-related areas, also showed that the fMRI signal is sensitive to changes in the object’s position. The response to a large central object shown in different places on the screen (changing by ~6 degrees) was greater than a repetition of the same image in the same position, indicating a partial recovery from adaptation. Our signal showed a complete recovery from adaptation, probably because the changes in our paradigm were greater (~8 degrees), involving crossing the vertical midline, and our objects were smaller and presented in the periphery. This is also consistent with the results of the findings of DiCarlo and Maunsell (2003).

Modulation by Eye Position

None of the previous studies were suitable to study the coordinate frame of neurons in LOC because the translation of the object was both on the retina and relative to the head. In this study (experiment 2), we were able to address this issue by changing the objects’ retinal position and head position independently. Our results show that the representation of some neurons in LOC is beyond a strictly retinotopic one and seems to be (at least) in head-based coordinates. Such a representation can be established if information about the eye position reaches LOC.

Recent studies indicate that modulation of neuronal activity by eye position can be found both in the monkey (Dobbins and others 1998; Bremmer 2000; Nowicka and Ringo 2000) and in human (DeSouza and others 2002) ventral areas. In area V4 of the monkey, about 50% of the neurons show gaze-dependent activation (Bremmer 2000). These cells have receptive fields in retinotopic coordinates but modulate their response according to the direction of gaze (having planar gain fields). Similar activity was found in monkey’s IT cortex (Nowicka and Ringo 2000). This indicates that there is a gradual change in the coordinate frame along the ventral stream. Although typically the receptive fields of neurons in V1 are retinotopic, higher areas along the pathway are influenced also by the eye position (multiplexed with the retinotopic information, i.e., area V4, IT).

Dorsal Areas and Coordinate Transformation

Coordinate transformation in visual areas is usually mentioned in the context of the dorsal stream in monkeys (Duhamel and others 1997; Andersen and Buneo 2002; Pouget and others 2002). A recent fMRI study reported a similar remapping signal in homologous areas in humans (Merriam and others 2003). Our experiment was not designed to reveal the transformation in these areas for 2 reasons. First, we rely on a localizer that is suited for the ventral pathway because it contains pictures of objects that are rather poor stimuli for the dorsal regions. These areas are much more active during presentation of dynamic scenes (movies) showing manipulation of such objects by the hands (Shmuelof and Zohary 2005). Second, the task we employed was a ventral task (requiring covert naming) rather than grasping or reaching for the object that would be more suitable for activating the dorsal pathway. As a result, only few subjects had active voxel clusters in the dorsal regions, and the degree of spatial consistency of these voxels between subjects was poor. This precluded the possibility of testing coordinate transformation in the dorsal areas in our study.

Extraretinal Representation in the Single-Neuron Scale

The simplest explanation in the single-neuron scale for the extraretinal coordinate frame could be the existence of a head-based receptive field. Neurons in LOC may possess a receptive field that covers a visual angle relative to the head and is not affected by gaze, similar to the neuronal receptive fields found in the monkey parietal cortex (Duhamel and others 1997). This explanation can account for both the adaptation and the contralateral effect.

One should note that retinotopic receptive field neurons with gain field modulation can account for the head-based contralateral effect but not for the adaptation. This option can still be relevant if the head-based adaptation is being received as an input from higher areas.

Irrespective of what the mechanism that incorporates the eye position is, it is clear that an extraretinal coordinate frame is utilized by significant population of neurons in LOC.

Possible Coordinate Frames

Our data suggest that the coordinate frame of representation of neuronal populations in LOC is beyond a strictly retinal one. It does not specify, however, what is the exact extraretinal representation being used by the neurons. Besides a head-centered coordinate frame, our results are also consistent with an egocentric (body based) representation. Because for obvious reasons we could not manipulate the subject’s head position relative to their body during the scan, we cannot rule out this alternative. Furthermore, an allocentric representation or an object-centered reference frame (assuming the screen is the object of reference) can also explain our results.

To summarize, we have evidence from 2 different functional imaging methodologies that the representation in the ventral object-related areas (LOC) is beyond a strictly retinotopic one. At present, we cannot tell whether this representation is in head, body, or allocentric coordinates. To discriminate between these possibilities, future experiments will require manipulating the head (or body) position while maintaining the same position of the object in the world. Ultimately, single-unit recordings from behaving monkeys are required to validate these propositions. In hindsight, it seems reasonable that neurons in object-related areas would represent the object’s position irrespective of the gaze direction because we often use saccadic eye movements when examining complex objects. For example, when viewing a face, observers typically focus on conspicuous and defining features (Yarbus 1967), such as the eyes, nose, mouth, or hairline. A head-based coordinate framework seems a useful neuronal property to establish a coherent image of the face across such saccades.

Supplementary Material

Supplementary material can be found at: http://www.cercor.oxfordjournals.org/.

Notes

We thank L. Daeuell, I. Rasmussen, I. Rabinowitch, and S. Geva for insightful comments. We also thank T. Orlov for the help with the 3D-cortex reconstruction and S. Lein for help with stimuli preparation. This study was funded by the Israel Science Foundation of the Israel Academy of Sciences grant #8009. Conflict of Interest: None declared. Address correspondence to email: udiz@lobster.huji.ac.il.

References

The Spatial Representation of Objects in LOC


