

# Cortical Representation of Medial Axis Structure

Mark D. Lescroart<sup>1</sup> and Irving Biederman<sup>1,2</sup>

<sup>1</sup>Neuroscience Graduate Program, University of Southern California, Los Angeles, CA 90089-2520, USA and

<sup>2</sup>Psychology Department, University of Southern California, Los Angeles, CA 90089-2520, USA

Address correspondence to Mark D. Lescroart, Hedco Neuroscience Building, 3641 Watt Way, Los Angeles, CA 90089-2520, USA.

Email: mark.lescroart@gmail.com.

**The identity of an object is not only specified by its parts but also by the relations among the parts. Rearranging parts can produce a completely different object, in the same manner as rearranging the phonemes in “fur” can yield “rough.” How does the visual system represent the relative positions of parts? Between-part relations can be characterized by specifying the relations between the medial axes (imaginary lines through the centers) of an object’s parts. A functional magnetic resonance imaging multivoxel classification study tested whether the medial axis structure is represented in the human visual system independent of part identity and overall object orientation. Stimuli were line drawings of novel 3-part geometrical objects, which differed in the relations between their parts’ medial axes (i.e., in their medial axis structures), the geons that composed each object, and the objects’ orientations in plane and in depth. In regions of interest throughout visual cortex, a support vector machine classifier was trained to distinguish objects that shared either the same medial axis structures or the same orientations. By the level of V3, different medial axis structures were more accurately classified than different orientations, indicating a change in the representation of shape compared with earlier visual areas.**

**Keywords:** fMRI, MVPA, object recognition, shape recognition, V3

## Introduction

Objects are represented as an arrangement of parts. Support for a parts-based representation derives from studies of behavior (Tversky and Hemenway 1984; Biederman and Cooper 1991; Biederman and Gerhardstein 1993; Hayward 1998), single unit electrophysiology (Tsunoda et al. 2001; Pasupathy and Connor 2002; Yamane et al. 2006), and neuroimaging (Hayworth and Biederman 2006). A critical challenge in the study of object representation is to determine how the relative positions of object parts are encoded. Rearranging parts can lead to a completely different interpretation of an object (Biederman 1987), just as changing the relative positions of phonemes in a word can change the meaning of the word (as in “rough” and “fur”). Explicit encoding of relations between parts is necessary to reason about object structure (Hummel and Biederman 1992; Hummel and Holyoak 2003) and to determine what parts of an object are missing, a task that appears on an IQ test for children (Wechsler 2004). Still, as essential as between-part relationships are to our understanding of the visual world, comparatively few studies have investigated how they might be encoded.

One way to define relationships between-object parts is in terms of the relative positions of the parts’ medial axes—the skeletal lines running through each part, as bones through fingers. More than 40 years ago, Harold Blum (Blum 1967;

Blum and Nagel 1978) observed that specifying an object’s medial axes provides a compact and intuitive way to parse the object into parts and thereby describe its structure. Many influential theories of object representation have used the concept of principal or medial axes to define the origin of an object-centered coordinate system (e.g., Marr and Nishihara 1978), to divide an object into parts (Hoffman and Singh 1997), or to define categorical relationships between parts (Biederman 1987). Recently, numerous variants of Blum’s Medial Axis Transform have been developed to reliably compute “shape skeletons” for 2D and 3D shapes (Dey and Sun 2006; Feldman and Singh 2006; Cornea et al. 2007), some of which have been suggested as a means to index online libraries of 3D graphical models (see <http://www.cs.princeton.edu/gfx/proj/shape/>).

Only a few neurocomputational studies have followed up on the broad and intuitive appeal of medial axes as shape descriptors. Lee et al. (1998) found that V1 cells show heightened responses to oriented bars located along the medial axis of a texture-defined figure, and Kimia (2003) has noted that the lateral connections in V1 are well situated to compute convex parts’ medial axes via a computation like Blum’s “grassfire” algorithm. To date, there has been no electrophysiological or imaging work exploring the representation of medial axes beyond V1.

Early computation of individual parts’ medial axes could lead to encoding of junctions between medial axes at later stages, analogous to the way that computation of local orientations in V1 is followed by encoding of junctions of edges (corners and curves) in V4 (Pasupathy and Connor 1999). In this study, we used multivoxel pattern analysis (MVPA) to test whether categorically different medial axis structures elicit reliably different blood oxygen level-dependent (BOLD) functional magnetic resonance imaging (fMRI) patterns in regions of interest (ROIs) throughout generally accepted cortical visual areas, using a set of novel objects that vary in their overall orientation, the shape of their parts, and their medial axis structures.

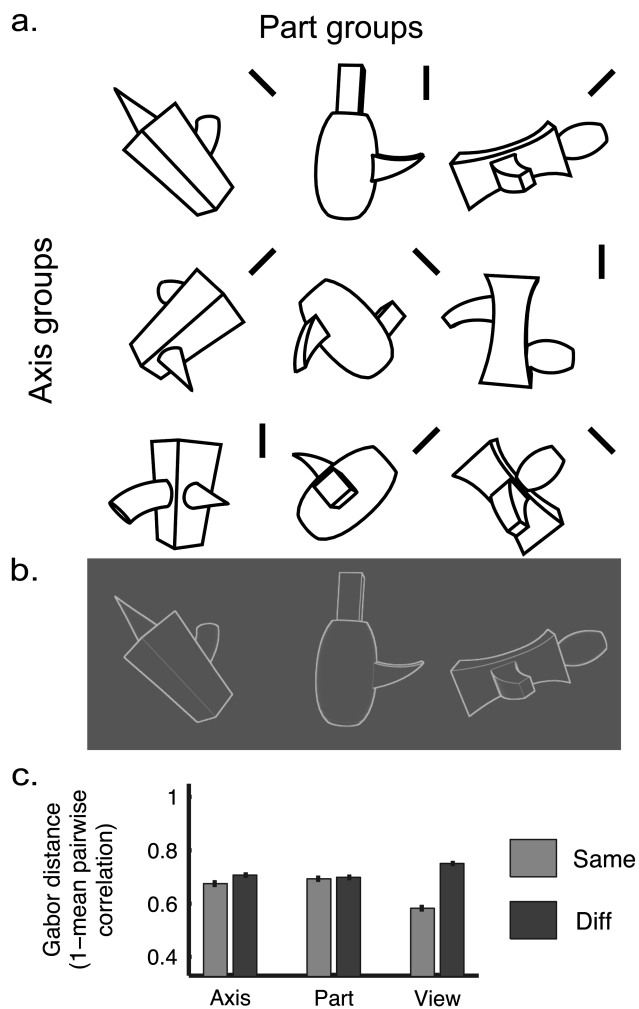
## Materials and Methods

### Subjects

Eight right-handed subjects (ages 21–29, 2 females) with normal or corrected-to-normal vision participated in the experiment. All were screened for safety and gave written informed consent before participating. They were financially compensated for their time, and all subject protocols were approved by the USC Institutional Review Board guidelines (and adhered to the Declaration of Helsinki).

### Stimuli

Our stimulus set consisted of 9 objects, each rendered from 6 different views (Fig. 1a). All objects were rendered in white on a dark gray



**Figure 1.** (a) Nine representative images (of the 54 images in the stimulus set). Each row shares the same medial axis structure (“axis groups”); each column shares the same component parts (“part groups”). View groups are marked by oriented bars (near vertical, tilted right, and tilted left). Bars were not displayed to subjects. (b) Stimuli as they appeared to the subjects, in contrast-equalized off-white on a dark gray background. (c) Average Gabor-jet distance between all pairs of stimuli within/between each group.

background with no shading or texture (Fig. 1*b*). The 9 objects were each composed of 1 of 3 groups of 3 geometrical volumes (geons), arranged in 1 of 3 different structures according to the relationships between the parts’ medial axes. The parts’ medial axes were conjoined according to categorical distinctions in medial axis relationships suggested in Biederman (1987), either end-to-end (i.e., with the medial axes of each part colinear) or end-to-side (i.e., with the medial axes of each part perpendicular). The parts joined end-to-side were either centered or offset and the 2 parts adjoining a larger part were either coplanar or offset.

To dissociate axis structure from low-level features such as local orientation and low-frequency outline, the overall orientation of the objects in plane and in depth was varied in six 22.5° increments. To assure that the variation in orientation did indeed change the low-level features of the images, stimuli were analyzed using a simple computational model of V1 (Lades et al. 1993). The model computed a “jet” of Gabor coefficients at each of 100 points arranged in expanding radial circles on each image. Each jet was composed of 40 Gabor filters: 8 equally spaced orientations (22.5° differences in angle) at 5 spatial scales, each centered on the same point in the image. The output of each filter was the magnitude of sine and cosine phases of spatial frequency at each location. The overall result for each image was a 4000-element vector (40 jets × 100 locations) that captured the local

orientation information in the same way that V1 theoretically does. A highly similar Gabor wavelet model can predict >30% of the variance in responses to natural images in V1 (more variance than is predicted by any other model) (David et al. 2004; Kay et al. 2008).

The low-level feature difference between each pair of images in our stimulus set was computed as one minus the Pearson correlation between the Gabor-jet vectors for each image. The average distances between images that either shared or did not share the same axis structure or overall orientation are shown in Figure 1*c*. The images that shared the same global orientation were more self-similar as a group by the Gabor-jet measure than were the images that shared the same axis structure. The Gabor-jet metric has been extensively used for scaling the physical differences between metrically varying stimuli (Fiser et al. 1996; Biederman and Kalocsai 1997; Xu et al. 2009) and predicts, almost perfectly, the psychophysical similarity of metrically varying faces and complex blobs (Yue et al. 2007).

The stimuli were thus designed such that the medial axis relationships between the objects’ parts were the only commonality among all the members of each “axis group.” Each image subtended ~5.8° of visual angle. All stimuli were generated using Blender (www.blender.org) and presented using the Psychophysics Toolbox (Brainard 1997; Pelli 1997; Kleiner et al. 2007) for Matlab (Mathworks).

#### Task: Attend to Component Parts

During the MRI scans, subjects attended to the identities of the geons composing the shapes and indicated via button press which of 3 part groups or families (columns in Fig. 1*a*) each shape belonged to. The shapes in the first group all had a straight-sided tapered brick as the central piece, with a cone and a curved cylinder attached to it. The shapes in the second group all had a large convex cylinder, a smaller straight-sided brick, and a smaller curved triangular prism, and the shapes in the third group all had a large concave brick, a smaller convex cylinder, and a smaller curved, tapered brick. Since each axis group and body orientation group contained an equal number of members of each part group, the task was orthogonal to the experimental manipulations of interest. Subjects used only one hand for their responses (half used their right hand, half their left).

In separate testing sessions, each subject also performed an analogous task identifying each axis structure group (rows in Fig. 1*a*) by button press in the same manner.

#### fMRI Data Collection and Preprocessing

MRI scanning was performed at USC’s Dana and David Dornsife Cognitive Neuroscience Imaging Center on a Siemens Trio 3-T scanner using a 12-channel head coil.  $T_1$ -weighted structural scans were performed on each subject using a magnetization-prepared rapid gradient echo (MPRAGE) sequence (TR = 1950 ms, TE = 2.26 ms, 160 sagittal slices, 256 × 256 matrix size, 1 × 1 × 1 mm voxels). Functional images were acquired using an echo planar imaging pulse sequence (TR = 2000 ms, TE = 30 ms, flip angle = 65°, in-plane resolution 2 × 2 mm, 2.0 or 2.5 mm-thick slices, 31 roughly axial slices). Slices covered as much of the brain as possible, though often the temporal poles and the crown of the head near the central sulcus were not scanned (due to large head size).

Subjects were scanned in 7 or 8 scanning runs of 55 trials each. Each trial consisted of a single stimulus presentation for 200 ms, followed by a 7.8 s fixation. Stimuli were presented in pseudorandom order (counterbalanced for axis groups).

fMRI data were collected using PACE online motion correction (Thesen et al. 2000). Additionally, data were temporally interpolated to align each slice with the first slice acquired, motion corrected (trilinear-sinc interpolation), and temporally smoothed to remove low-frequency drift (kernel = 3 cycles/run). All preprocessing was carried out using Brain Voyager QX version 2.08 (Brain Innovation, Maastricht, the Netherlands) (Goebel et al. 2006). Data were not smoothed or normalized; ROIs were transformed to the functional data’s space, and all pattern analysis was done in native functional space. The raw activation values for time points from 4 to 6 s after stimulus onset (2 sequential TRs worth of data) on each trial were averaged to create a single activity value per trial. All trial values were

converted to  $z$  scores (by run) prior to classification analysis to minimize baseline differences between runs.

Because each trial consisted of only a single presentation of an image (rather than a block of different images of the same class), it was possible to relabel trials and attempt to classify different groups within the same data set. Thus, we were able to compare how well a given region distinguished objects with different axis structures and compare that with how well the same region distinguished different orientations of the composite objects, using the same data.

### Regions of Interest

ROIs (Fig. 2a) were defined using independent localizer scans and anatomical criteria. Rotating contrast-reversing wedges were used to define V1–V4 and V3A (as in Engel et al. 1997; Sereno 1998). Wedges spanned 8.9 visual degrees from center to periphery and 45 radial degrees. Lateral occipital cortex (LO) was defined as the region more active to objects than scrambled versions of the same objects ( $t$  contrast with the false detection rate set at  $P < 0.05$ ), spanning the region from the dorsal part of V3 (dorsally) to V4 (ventrally) (Grill-Spector et al. 1999). We also defined a ventral visual region encompassing the fusiform face area, the parahippocampal place area, and shape-selective regions in the posterior fusiform gyrus (pFs) by a contrast of faces + scenes + objects > scrambled objects. (These regions were initially analyzed separately, but no differences were found, so they were grouped together for simplicity.) Stimuli for the object/face/place localizer subtended  $\sim 6^\circ$  of visual angle (approximately the same size as the images in the main experiment). A region in the intraparietal sulcus (IPS) was defined by mixed anatomical and functional criteria: we took the region extending dorsally up the medial bank of the IPS from V3A to a region that showed increasing activation to increasing working memory load (as in Xu and Chun 2006). Finally, in 5 of the 8 subjects, unilateral ROIs in the right and left motor cortex were defined along the anterior banks of the central sulcus. (In the other 3 subjects, our scanning protocol covered less than 50% of the motor cortex due to larger head sizes, so no ROIs were defined.)

Since the ROIs varied substantially in size and mean activation level, both of which have been shown to influence classification performance

(Cox and Savoy 2003; Smith et al. 2010), we imposed 2 further restrictions on each region. First, for each ROI, we sorted the voxels according to their overall responsiveness ( $t$  statistic) to all axis groups and chose only voxels that showed a significant ( $t > 2$ ,  $P < 0.05$  uncorrected) response to a contrast of all stimulus conditions versus fixation (in the training runs only). Second, we chose only the 300 most-responsive voxels in each region to keep the number of voxels constant across all ROIs.

### fMRI Classification Analyses

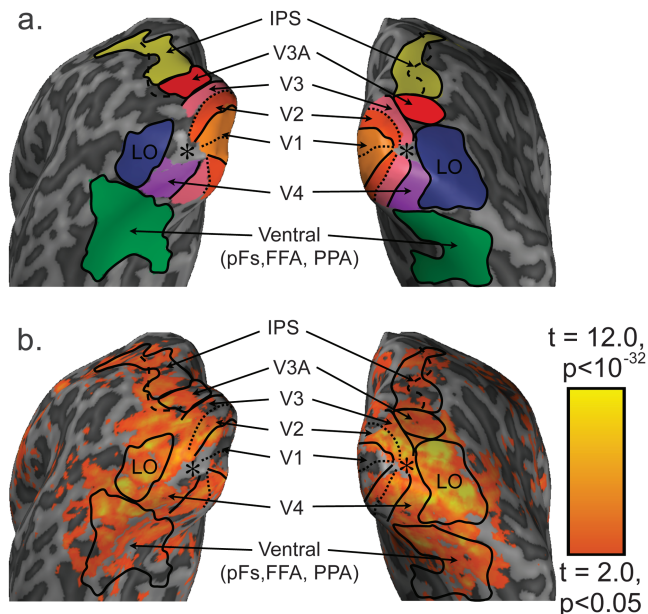
We used a linear support vector machine (SVM) classifier to assess whether the 3 axis groups elicited reliably different patterns of activation in each ROI. Linear SVMs have been widely used in fMRI multivoxel pattern classification studies (Eger et al. 2008; Ester et al. 2009; e.g., Kamitani and Tong 2005; Ostwald et al. 2008) and have been shown to be more sensitive at detecting pattern differences than other multivariate measures (Cox and Savoy 2003). Our SVM classifier was implemented via the Python Multivariate Pattern Analysis package (Hanke et al. 2009; www.pympva.org) using the LibSVM library. The soft margin parameter ( $c$ ) was scaled for each subject and ROI by dividing by the square root of the norm of the data. The SVM classifier was trained on all but one of the fMRI runs and tested on the withheld run. Each of the runs were withheld as the test set once in an  $n$ -fold cross-validation, for a total of 440 test trials in subjects with 8 runs and 385 trials in the 1 subject with 7 runs.

## Results

### Behavioral Results

Subjects were readily able to assign each object to its appropriate “part group.” Mean accuracy was 98.1% correct (essentially at ceiling) and mean reaction time (RT) was 751 ms, with no reliable differences across experimental runs in RTs or error rates (repeated measures analysis of variance [ANOVA], both  $F_{7,7} < 1.2$ ,  $P > 0.30$ ). Nor were there any reliable differences between part groups, orientations, or axis structures in either RTs or error rates. It should be noted that although subjects were making judgments about objects’ parts, there was a trend toward differences in RTs for objects belonging to different axis families, most likely because of greater self-occlusion between parts in the third axis family in several of the views (Fig. 1a, third row), which made part judgments slightly more difficult. (For RT differences in judging part families,  $F_{7,2} = 3.27$ ,  $P = 0.07$ ; all other  $F_s < 1.75$ ,  $P > 0.13$ .)

In the complementary task (conducted in separate sessions), the same subjects were also highly accurate (98.6%) in assigning each object to its axis group with mean RTs of 794 ms, again with no indication of improvement across runs (after training) in either RTs or error rates (both  $F_{7,7} < 1.1$ ,  $P > 0.40$ ). Subjects showed immediate understanding of the task with near ceiling performance. They identified the first medial axis family (Fig. 1a, row 1) more quickly than the other two: mean RT of 743 ms for that family versus 812 and 827 ms for the other two,  $F_{7,2} = 6.03$ ,  $P = 0.013$ , post hoc test (Tukey’s honestly significant difference [HSD]) for axis family 1 versus both 2 and 3,  $P < 0.05$ . This advantage for the first family was likely due to its distinctive elongation relative to the other structures. Unlike the part-group task, subjects were also slower at judging the axis structure of the stimuli rotated the farthest from vertical: for the most extreme orientations mean RT = 835 ms; for vertical, 785 ms;  $F_{7,5} = 10.16$ ,  $P < 0.0001$ ; Tukey’s HSD post hoc test comparing vertical with extreme orientations,  $P < 0.05$ . All 3 axis groups—even the first group (Fig. 1a, row 1) which, as noted previously, appeared distinctive



**Figure 2.** (a) ROIs for a representative subject, displayed on a posterior view of an inflated brain. ROIs were defined using independent localizers and anatomical criteria. Dotted lines represent the horizontal meridian, solid lines represent the vertical meridian, asterisks represent the foveal confluence in each hemisphere, and the dashed line marks the IPS. The ventral region contained face- and place-selective voxels as well as object-selective voxels. Motor cortex ROI not shown. (b) Activation maps of response to all stimuli ( $t$  values for contrast of all conditions–fixation).



from the other two—showed significant costs of recognition (greater RTs) at the orientations farthest from vertical.

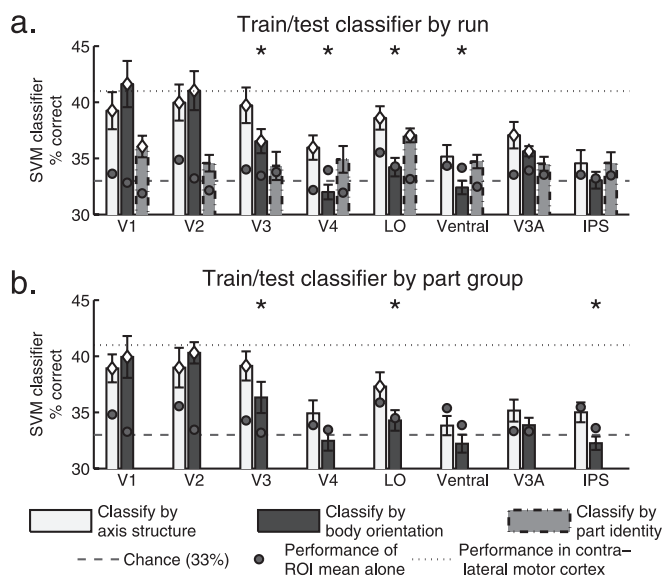
### Univariate fMRI Results

We saw activation throughout generally accepted visual areas (Fig. 2*b*) in response to all of our conditions, with the most (and most significant) activation in the lateral occipital cortex and surrounding regions. (For BOLD response curves for each region, see Supplementary Results.)

### fMRI Classification Results

All regions from V1 to LO were able to distinguish the 3 different axis structures (i.e., the 3 different arrangements of the objects' parts) significantly better than chance (all  $t_{s7} > 2.43$ ,  $P < 0.05$ ) (Fig. 3*a*). In V1 and V2, the classifier performed slightly better at distinguishing different orientations of the objects (although this difference was not significant). By the level of V3, however, significantly more accurate classification was obtained for distinctions between medial axis structures than for distinctions between body orientations:  $t_7 = 2.87$ ,  $P = 0.02$ . In the ventral and parietal ROIs, the same trend was observed, though overall classification performance did not exceed chance: both  $t_{s7} < 1.90$ ,  $P > 0.10$ . (See Supplementary Table 1 for P values for all statistical tests. All t-tests are two-tailed paired t-tests.)

To assess whether there was an interaction between stage in the visual hierarchy and classification accuracy for axis structure and orientation, we ran a 2-way repeated measures ANOVA, with factors ROI (5 levels: V1, V2, V3, V4, and LO) and CLASSIFIER TASK (2 levels: classify by axis structure, classify by orientation). There was a significant interaction between ROI and CLASSIFIER TASK,  $F_{4,28} = 6.53$ ,  $P < 0.001$ .



**Figure 3.** SVM classifier results by ROI. (a) Mean classifier accuracy when classifier was trained on all but one of the MRI scans and tested on the last scan. The dotted line around the bar for classification by part identity indicates that the classifier task matched with the subjects' behavioral task. (b) Mean classifier accuracy when the classifier was trained on 2 of the part families and tested on the third (test of generalization to new stimuli sharing the same axis structure). Asterisks indicate significant differences between axis structure and body orientation classification, the upper dotted line shows classification accuracy in contralateral motor cortex, and white diamonds at the bars' peaks indicate significantly better-than-chance classification:  $t_7 > 2.43$ ,  $P < 0.05$ . Error bars are standard error of the mean.

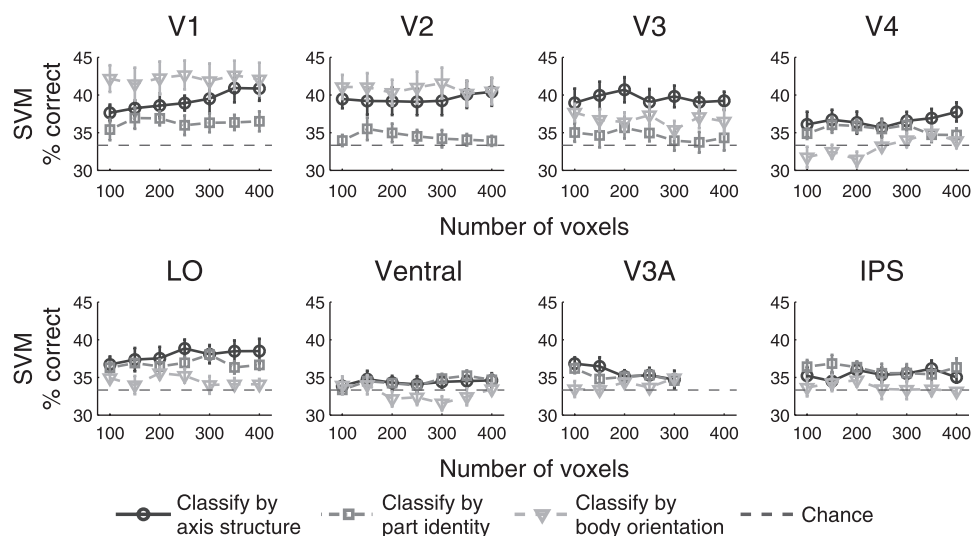
A similar pattern of results was observed if we used exactly the same number of voxels in each ROI (from 50 to 400 voxels; Fig. 4) as well as if we used all voxels within each ROI. Note that with fewer voxels fed to the classifier, V1 classified orientation substantially more accurately than axis structure. For 100-, 200-, and 250-voxel patterns, this difference was significant in V1,  $t_7 > 2.7$ ,  $P < 0.05$ .

In order to assure that each axis family could be distinguished from both other axis families, we plotted the confusion matrices of classifier responses. Confusion matrices for V3, V4 and LO (regions for which axis structure classification exceeded orientation classification) are shown in Figure 5. In V3, all groups could be distinguished above chance (all  $t_s > 2.5$ ,  $P < 0.05$ ) and in LO, 2 of the 3 groups could be distinguished from the others ( $t_s > 2.8$ ,  $P < 0.05$ ). For the third group (axis family 2), the correct group was chosen most often, but the classification accuracy fell short of significance,  $t = 1.92$ ,  $P = 0.09$ . In V4, none of the groups could be distinguished from chance individually ( $t \leq 2.0$ ,  $P > 0.08$ ).

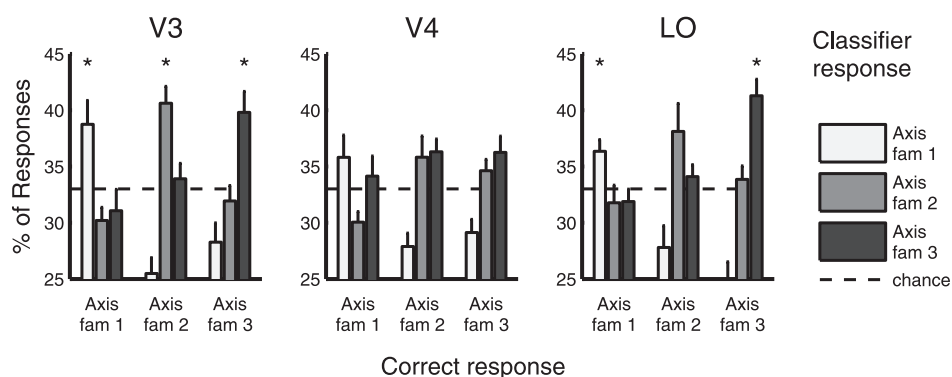
Even though the subjects were making explicit judgments on each trial as to which part group each image belonged to (and thus presumably attending to the features that distinguished the different part groups), in none of the ROIs were part groups more accurately classified than the axis structure groups. Classification by parts was significantly more accurate than chance in V1,  $t_7 = 2.85$ ,  $P = 0.02$  and LO,  $t_7 = 6.03$ ,  $P < 0.0001$ . There was only a trend toward classification by parts in LO being better than classification by orientation,  $t_7 = 2.15$ ,  $P = 0.067$ . The interpretation of the higher accuracy for part-group classification in LO is complicated by the congruence with the subjects' task. Nonetheless, the higher classification accuracy in LO is noteworthy, particularly given the lack of sensitivity shown by V2-V4 to the parts (vs. orientation).

For the 5 subjects for whom we had data from the motor cortex, mean classification accuracy in the ROI contralateral to the hand each subject used for his or her response was 41.0% for the part groups versus 32.6% for axis groups and 32.0% for orientation groups. Accuracy on the side ipsilateral to the response hand was 35.9% for part groups, 31.8% for axis groups, and 33.6% for view groups. This pattern of accuracy serves as a sanity check (accurate classification of parts was to be expected, given that subjects were making one-handed responses to the part groups). It is also interesting to note that classification accuracy for part families was approximately equal to the accuracy observed in the visual regions for classification by views or axis structures (Fig. 3); however, statistically the accuracy for part classification fell short of significance, most likely due to the limited number of subjects (5 instead of 8);  $t_4 = 2.65$ ,  $P = 0.057$ .

Since the classifier was tested on novel instances (trials) of each of the stimuli, and not completely novel stimuli, it is possible that the voxels in each ROI (and thus the classification algorithm) could have picked up on some idiosyncratic feature of each axis structure group rather than axis structure per se. For example, cells in macaque posterior inferotemporal cortex have been shown to respond to particular combinations of adjacent boundary curves (Brincat and Connor 2004), such as those that might occur at the junction between 2 parts in our stimuli. For a more rigorous test of whether these regions represented axis structure and not more local features, we trained the SVM classifier on trials of 2 of the 3 part groups and tested it on the third (each part group was left out in turn



**Figure 4.** Classification accuracy for equivalent numbers of voxels in each ROI. Error bars are standard error of the mean.



**Figure 5.** Classifier confusion matrices for V3–LO. Asterisks indicate significantly above-chance performance (two-tailed  $t$  test,  $P < 0.05$ ). Error bars are standard errors of the mean.

a 3-fold cross-validation). Note that we have specifically chosen parts that varied in dimensions (e.g., curvature/pointedness, convexity/concavity) that have been shown to modulate neural activity in both human lateral occipital cortex (Op de Beeck et al. 2008) and macaque inferotemporal cortex (Kayaert et al. 2003, 2005) and V4 (Pasupathy and Connor 1999), thus making it less likely that objects with different parts will elicit similar patterns of activation. Nonetheless, even when tested on stimuli composed of different parts than the stimuli in the training data, the classifier based on voxels in V3 and LO still distinguished different axis structures above chance and better than different body orientations (Fig. 3*b*, for all  $t$  and  $P$  values, see Supplementary Table 1). Classification performance was slightly lower overall than when trained and tested by runs, but the classifier was also trained on fewer trials (2/3 of the data set vs. 7/8 for training and testing by runs).

It is worth noting that there is a slight risk of circularity in this analysis compared with our main analysis. In our main analysis, SVM training and testing were performed on separate scanning sessions, and voxel selection was performed based only on the training sessions. In this analysis, the training and testing data were drawn from interleaved trials in the same scanning sessions, and voxel selection was performed based on a main-effect analysis spanning all the scans. We call the risk

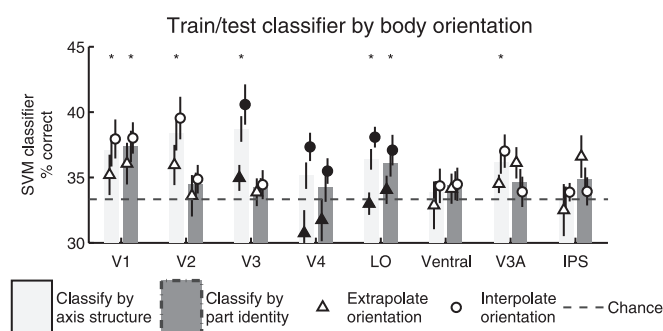
“slight” because, for our data, selection based solely on the training set chose  $91.3 \pm 1.2\%$  of the same voxels as selection based on the whole data set (averaged across subjects, runs, and ROIs). Furthermore, more than 99% of the 200 most-responsive voxels were chosen by both methods of voxel selection. In other words, the voxels that were (perhaps spuriously) selected by “bad” method but not the “good” method constituted a small minority (~8.7%) of the total number of voxels and had smaller response magnitudes than at least 2/3 of the voxels in each ROI. Thus, the different methods for voxel selection were unlikely to have had a strong effect on classification accuracy.

It is possible that some statistical dependence could exist between pairs of the conditions due to training and testing on interleaved trials, but we find that highly unlikely as well. First, the trials were widely spaced (8 s apart) and counterbalanced such that each axis group appeared before every other an equal number of times, making it highly unlikely that trials for one axis group were systematically biased by interaction with other axis groups. Second, we still observed poor classification results in some regions (in V3A, V4, ventral, and IPS regions), indicating that whatever dependence there might have been between the training and testing sets, that dependence was not sufficient to explain the above-chance classification.

Furthermore, our most critical measure is a comparison between 2 classification schemes (classification by common axis structure and by common body orientation), both of which should have benefited equally from any statistical dependence between the training and testing sets—and yet the advantage of classification by axis structure over classification by body orientation remained.

We performed a similar test to see whether accurate classification of medial axis structures could be achieved over different views of the objects: we trained the classifier on 5 of the views of each object and tested it on the sixth. Each orientation was left out as the testing set once in successive cross-validation steps. Overall, classification accuracy for axis structure groups was above chance for V1–V3, V3A, and LO,  $t_7 > 3.38$ ,  $P < 0.05$  (Fig. 6). For a more rigorous test of whether axis structure groups elicited consistent patterns over different views, we separated out the different cross-validation splits of the data and recombined them in 2 ways. First, we took the average accuracy for cross-validation splits for which the extreme orientations (ca.  $-45^\circ$  and ca.  $+67.5^\circ$ ) were left out as the testing set—that is, the data sets for which the classifier had to extrapolate to a novel orientation. Second, we took the average accuracy for splits in which one of the intermediate orientations (ca.  $-22.5^\circ$  to ca.  $+45^\circ$ ) was left out as the testing set—that is, data sets for which the classifier could interpolate to a novel orientation. For V3, V4, and LO (all regions showing an increased sensitivity to axis structure vs. body orientation), classification accuracy was significantly better in trials for which the classifier could interpolate:  $t_7 > 2.7$ ,  $P < 0.05$  (Fig. 6). The only reversal of this trend was in the parietal lobe, for classification by part families (which matched with the subjects' task), although this trend did not reach significance:  $t_7 = 1.42$ ,  $P = 0.20$ .

Because the overall classification accuracy was relatively low compared with other MVPA classification studies, we used 2 additional nonparametric measures—bootstrapping random assignments of trial labels and group assignments—to determine whether classification accuracy for axis structure groups was significantly better than chance (see Supplementary Methods). These more conservative tests also confirmed the statistical significance of the results (see Supplementary Table 1).



**Figure 6.** Classification accuracy when the classifier was trained and tested on stimuli with different body orientations. Bars represent average classification accuracy across all splits of the data and asterisks indicate classification accuracy significantly greater than chance ( $P < 0.05$ ). Filled markers (triangles and circles) indicate significant difference ( $P < 0.05$ ) between interpolation and extrapolation splits. The dotted line around the bar for classification by part identity represents that the classifier task matched with the subjects' behavioral task.

Since a SVM is sensitive to even small differences in mean activation, above-chance classification in the range that we observed ( $\sim 36\%$  to  $39\%$ ) could potentially be achieved even using a one-dimensional measure such as the mean activity if a simple threshold would suffice to distinguish one group from the others for a sufficient number of trials. Thus, the classification analysis was repeated using only the mean activity for each ROI instead of the full pattern of voxel activity in each ROI (as in Meyer et al. 2010). All regions from V1 to LO showed greater classification accuracy when the voxel patterns were used compared with when the mean was used (see circles in Fig. 3; for statistical values, see Supplementary Table 1), indicating that the information about axis structure was present in the spatial profile of activation rather than simply the average activation of each region.

Because some of the subjects had training on the axis structure stimuli before the scanning session, another possibility is that the accurate classification of axis structures was a result of learning rather than a stimulus-driven effect. Training on stimulus classes (though typically over several sessions) has been shown to change BOLD fMRI responses in shape-selective areas (Op de Beeck et al. 2006; Yue et al. 2006). To test for an effect of learning, we split our subjects into 2 groups—those who had performed the axis structure behavioral task before the scan, and those who had not—and ran a 2-way repeated measures ANOVA, with factors TASK ORDER (axis first, parts first) and CLASSIFICATION ACCURACY (axis classification, orientation classification), for each of our ROIs. We found no main effect of TASK ORDER (all  $F_{s_{1,4}} < 1.8$ ,  $P > 0.25$ ) nor any interaction of TASK ORDER and CLASSIFICATION ACCURACY, all  $F_{s_{1,4}} < 0.7$ ,  $P > 0.40$ . We also found no relationship between multivoxel classification accuracy and RT variability between conditions nor between classification accuracy and trial-to-trial variation in RT (see Supplementary Results).

One aspect of behavior that did covary with classification accuracy was mean RT across conditions for the part judgment task performed during the fMRI data acquisition. Overall RT was negatively correlated with classification accuracy in LO ( $r = -0.79$ ,  $P < 0.05$ ). It is somewhat surprising that RTs in the part task, but not the axis task, should negatively correlate with classification accuracy. (Greater overall RTs in the part task would not mean more difficulty “processing” axis structure.) However, since the task was reported by all subjects to be extremely easy, long RTs may be reflective of boredom, fatigue, or other disengagement with the task and the stimuli. Disengagement, in turn, could likely result in less BOLD signal and lower classification accuracy.

## Discussion

MVPA revealed more accurate classification of objects with the same medial axis structure than objects with the same body orientation in intermediate visual areas, beginning in V3. This was not a low-level (retinotopic) effect, since V1 showed the opposite ordering of classification accuracy, with orientation  $>$  axis structure (Figs 3a and 4), and a simple computational model of V1 showed greater similarity among objects sharing the same orientation than objects sharing the same axis structure (Fig. 1c). V3's pattern of classification accuracy (axis structure  $>$  orientation) was maintained when the classifier was tested on stimuli not used in the training set (Fig. 3b),



indicating that V3 voxels are sensitive to arrangements of medial axes despite considerable variation in other dimensions.

Structural information present in V3, V4, and LO was still somewhat orientation dependent in that the SVM could not extrapolate to classify axis structures outside the range of trained orientations (Fig. 6). Rather than viewing this as a failure to achieve full view invariance, we suggest that encoding of relations between parts (at least at the level of V3) specifies gravitational relations such as “top-of” as well as axis structural relations. Indeed, rotating objects in plane incur costs in object identification (Jolicoeur 1985; Tarr and Pinker 1989; Hayward et al. 2006), so full 2D rotation invariance would not be an accurate characterization of human vision (Hummel and Biederman 1992).

### **Relation to Other Work**

Compared with V1 and V2, not much is known about V3. Most cells in macaque V3 show orientation tuning, sometimes with multi-peaked tuning curves (Felleman and Van Essen 1987; Anzai et al. 2007). Many V3 cells also show end stopping and binocular disparity tuning (Felleman and Van Essen 1987; Gegenfurtner et al. 1997). V3 receives direct inputs from V1 with major inputs from layer 4B, which is associated with the magnocellular pathway and processing of low spatial frequency information (Felleman et al. 1997). These results are compatible with a role for V3 in encoding medial axis structure (though the stimuli used in the cited studies were too simple for any effect of axis structure to be evident). V3 is arguably the last visual stage before the ventral and dorsal pathways diverge (Ungerleider and Mishkin 1982; Felleman and Van Essen 1991). The dorsal stream has been implicated in spatial reasoning (e.g., mental rotation tasks), whereas the ventral stream has been implicated in the recognition of objects despite variation in view (Gauthier et al. 2002; Vanrie et al. 2002; Wilson and Farah 2006). Since V3 projects both dorsally and ventrally, medial axis information computed by V3 could feed into both processes.

LO has been implicated in the processing of between-part relations in studies of patient SM by Behrmann and colleagues (Behrmann et al. 2006; Konen et al. 2011). SM, who had a lesion in ventral LO, had difficulty distinguishing objects differing only in the relations between their parts, despite a preserved ability to detect variations in part shape. SM’s lesion was clearly anterior to V3 (Konen et al. 2011), suggesting that structural computations in V3 may not be “read out” until the signal has reached LO. LO also shows strong sensitivity to between-object relations, independent of the objects’ absolute spatial positions (Kim and Biederman 2010; Hayworth et al. 2011).

### **Why Was the Difference in Classification Accuracy between Orientation and Axis Structure Not Greater in V1?**

Though classification accuracy for axis structure and orientation was comparable for larger voxel patterns, when fewer V1 voxels were used for classification (e.g., 100 voxels vs. 400 voxels, Fig. 4), V1 did classify orientation significantly more accurately than axis structure. Thus, the most-responsive voxels in V1 were indeed most selective for orientation. The comparable performance for classification by body orientation and axis structure with larger voxel patterns may be due to

a ceiling effect; no region in our study classified any parameter (axis structure, parts, or orientation) better than ~42%.

### **Why the Lower Accuracy for Classification of Parts versus Axis Structures?**

Voxels in LO can distinguish “pointy” objects from smoothly curved or blocky objects (Op de Beeck et al. 2006, 2008). However, all of the objects in the present investigation contained some blocky parts, some curved parts, and some pointy parts. Thus, classification of part groups was likely more difficult for lack of a single (nonaccidental) distinguishing shape attribute.

### **Why the Low Classification Accuracy Overall?**

The fMRI signal for single trials is much noisier than the signal for blocks of sequentially presented objects. However, our design depended critically on single trial presentations (so we could relabel trials to reflect different aspects of the stimuli). We thus sacrificed a degree of fMRI signal strength for theoretical clarity. In addition, to achieve control of stimulus features, the images we used were far more similar overall than stimuli used in many other multivoxel experiments (e.g., Haxby et al. 2001; Eger et al. 2008; Kriegeskorte et al. 2008), which differed in color, texture, and form, as well as semantic category, familiarity, behavioral utility, and evolutionary significance. Thus, classification accuracy for our stimuli might reasonably be expected to be lower since successful classification must depend on specific subtle differences in shape. A final reason for reduced accuracy is that different features are encoded by different neurons within single voxels. For example, different neurons in V4 may encode color or contour curvature (Zeki 1973; Pasupathy and Connor 1999, 2001). Responses to features other than axis structure—for example, local boundary contour curvature—would be manifested as noise in our experiment.

### **Interpretation of MVPA Results**

Given the certainty that multiple features are encoded by V3 neurons, how should above-chance classification accuracy for medial axis structure be interpreted? One possibility is that there are simply more neurons tuned for axis structure than for orientation in V3 and subsequent regions. However, fMRI signals are biased toward signals that are mapped across the cortex at the scale of fMRI voxels (Drucker and Aguirre 2009; Freeman et al. 2011). Thus, another plausible interpretation for our findings is that there is a change in the organization of the representation in V3 that favors readout of axis structure at the scale of fMRI. Many theorists have suggested that the cortex is organized to minimize wiring length for critical computations (Allman 1999; Cherniak et al. 2004; Chklovskii and Koulakov 2004). Thus, either interpretation—a change in the proportion of neurons encoding axis structure or a change in cortical organization (or a combination of both)—is consistent with a role for V3 in encoding medial axis structure. (For further discussion of the role of axis structure compared with other dimensions, see Supplementary Material.)

The lower classification accuracy in the ventral ROI (vs. LO) was somewhat surprising, given the known role for the posterior fusiform gyrus in encoding shape (Haxby et al. 2001; Kourtzi and Kanwisher 2001; Hayworth and Biederman 2006). However, several other studies have also found poorer

classification of novel objects in the posterior fusiform gyrus than in LO (Williams et al. 2007; Op de Beeck et al. 2008; Drucker and Aguirre 2009). Again, this may reflect a change in the organization of the region—there may still be neurons sensitive to axis structure that are not clustered sufficiently to differentially influence the BOLD signal in different voxels. Alternatively, several studies have suggested that regions in ventral temporal cortex respond to particular semantic categories (e.g., animals, body parts, faces) more than visual shape features per se (Kiani et al. 2007; Mahon et al. 2009).

## Conclusions

Our results demonstrate that information about the relative positions of an object's parts, characterized by its medial axis structure, is encoded at particular retinotopic (or gravitational) orientations in V3 and successive visual stages. Clearly, axis structure is not the only feature encoded by V3 or any of the other regions, nor does the entire world look like stick figures. But facile object classification is critically dependent on specification of the relations between parts—relations that are well defined by axis structure. Many of the categories of objects that have been shown to be represented in anterior ventral visual regions such as tools and animals differ greatly in their medial axis structures. Moreover, spatial abilities known to be mediated by the parietal lobe (such as mental rotation) may rely on computation of medial axis structures (Just and Carpenter 1976). Thus, a representation of medial axis structure in V3 could provide a link between local feature tuning in V1 and higher order processing in both the dorsal and the ventral visual pathways.

## Supplementary Material

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

## Funding

National Science Foundation Division of Behavioral and Cognitive Sciences (grants 04-20794, 05-31177, and 06-17699 to I.B.).

## Notes

We wish to thank Jonas Kaplan, Bosco Tjan, Kenneth Hayworth, Jiye Kim, Xiaokun Xu, and Ori Amir for advice, assistance, and useful discussions and Jiancheng Zhuang for his support of the scanner. *Conflict of Interest:* None declared.

## References

Allman JM. 1999. *Evolving brains*. New York: Freeman.  
 Anzai A, Peng X, Van Essen DC. 2007. Neurons in monkey visual area V2 encode combinations of orientations. *Nat Neurosci*. 10:1313–1321.  
 Behrmann M, Peterson MA, Moscovitch M, Suzuki S. 2006. Independent representation of parts and the relations between them: evidence from integrative agnosia. *J Exp Psychol Hum Percept Perform*. 32:1169–1184.  
 Biederman I. 1987. Recognition-by-components: a theory of human image understanding. *Psychol Rev*. 94:115–147.  
 Biederman I, Cooper EE. 1991. Priming contour-deleted images: evidence for intermediate representations in visual object recognition. *Cogn Psychol*. 23:393–419.

Biederman I, Gerhardstein PC. 1993. Recognizing depth-rotated objects: evidence and conditions for three-dimensional viewpoint invariance. *J Exp Psychol Hum Percept Perform*. 19:1162–1182.  
 Biederman I, Kalocsai P. 1997. Neurocomputational bases of object and face recognition. *Philos Trans R Soc Lond B Biol Sci*. 352:1203–1219.  
 Blum H. 1967. A transformation for extracting new descriptors of shape. In: Wathen-Dunn W, editor. *Models for the perception of speech and visual form*. Cambridge, MA: MIT Press. p. 362–380.  
 Blum H, Nagel RN. 1978. Shape description using weighted symmetric axis features. *Pattern Recogn*. 10:167–180.  
 Brainard DH. 1997. The Psychophysics Toolbox. *Spat Vis*. 10:433–436.  
 Brincat SL, Connor CE. 2004. Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nat Neurosci*. 7:880–886.  
 Cherniak C, Mokhtarzada Z, Rodriguez-Esteban R, Changizi K. 2004. Global optimization of cerebral cortex layout. *Proc Natl Acad Sci U S A*. 101:1081–1086.  
 Chklovskii DB, Koulakov AA. 2004. Maps in the brain: what can we learn from them? *Annu Rev Neurosci*. 27:369–392.  
 Cornea ND, Silver D, Min P. 2007. Curve-skeleton properties, applications, and algorithms. *IEEE Trans Vis Comput Graph*. 13:530–548.  
 Cox DD, Savoy RL. 2003. Functional magnetic resonance imaging (fMRI) “brain reading”: detecting and classifying distributed patterns of fMRI activity in human visual cortex. *Neuroimage*. 19:261–270.  
 David SV, Vinje WE, Gallant JL. 2004. Natural stimulus statistics alter the receptive field structure of v1 neurons. *J Neurosci*. 24:6991–7006.  
 Dey TK, Sun J. 2006. Defining and computing curve-skeletons with medial geodesic function. *Proceedings of the Fourth Eurographics Symposium on Geometry Processing*. Cagliari, Sardinia (Italy): Eurographics Association. p. 143–152.  
 Drucker DM, Aguirre GK. 2009. Different spatial scales of shape similarity representation in lateral and ventral LOC. *Cereb Cortex*. 19:2269–2280.  
 Eger E, Ashburner J, Haynes JD, Dolan RJ, Rees G. 2008. fMRI activity patterns in human LOC carry information about object exemplars within category. *J Cogn Neurosci*. 20:356–370.  
 Engel SA, Glover GH, Wandell BA. 1997. Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cereb Cortex*. 7:181–192.  
 Ester EF, Serences JT, Awh E. 2009. Spatially global representations in human primary visual cortex during working memory maintenance. *J Neurosci*. 29:15258–15265.  
 Felleman DJ, Burkhalter A, Van Essen DC. 1997. Cortical connections of areas V3 and VP of macaque monkey extrastriate visual cortex. *J Comp Neurol*. 379:21–47.  
 Felleman DJ, Van Essen DC. 1987. Receptive field properties of neurons in area V3 of macaque monkey extrastriate cortex. *J Neurophysiol*. 57:889–920.  
 Felleman DJ, Van Essen DC. 1991. Distributed hierarchical processing in the primate cerebral cortex. *Cereb Cortex*. 1:1–47.  
 Feldman J, Singh M. 2006. Bayesian estimation of the shape skeleton. *Proc Natl Acad Sci U S A*. 103:18014–18019.  
 Fiser J, Biederman I, Cooper EE. 1996. To what extent can matching algorithms based on direct outputs of spatial filters account for human object recognition? *Spat Vis*. 10:237–271.  
 Freeman J, Brouwer GJ, Heeger DJ, Merriam EP. 2011. Orientation decoding depends on maps, not columns. *J Neurosci*. 31:4792–4804.  
 Gauthier I, Hayward WG, Tarr MJ, Anderson AW, Skudlarski P, Gore JC. 2002. BOLD activity during mental rotation and viewpoint-dependent object recognition. *Neuron*. 34:161–171.  
 Gegenfurtner KR, Kiper DC, Levitt JB. 1997. Functional properties of neurons in macaque area V3. *J Neurophysiol*. 77:1906–1923.  
 Goebel R, Esposito F, Formisano E. 2006. Analysis of functional image analysis contest (FIAC) data with Brainvoyager QX: from single-subject to cortically aligned group general linear model analysis and self-organizing group independent component analysis. *Hum Brain Mapp*. 27:392–401.  
 Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzhak Y, Malach R. 1999. Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron*. 24:187–203.



- Hanke M, Halchenko YO, Sederberg PB, Olivetti E, Frund I, Rieger JW, Herrmann CS, Haxby JV, Hanson SJ, Pollmann S. 2009. PyMVPA: a unifying approach to the analysis of neuroscientific data. *Front Neuroinform*. 3:3.
- Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, Pietrini P. 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*. 293:2425–2430.
- Hayward WG. 1998. Effects of outline shape in object recognition. *J Exp Psychol Hum Percept Perform*. 24:427–440.
- Hayward WG, Zhou G, Gauthier I, Harris IM. 2006. Dissociating viewpoint costs in mental rotation and object recognition. *Psychon Bull Rev*. 13:820–825.
- Hayworth KJ, Biederman I. 2006. Neural evidence for intermediate representations in object recognition. *Vision Res*. 46:4024–4031.
- Hayworth KJ, Lescroart MD, Biederman I. 2011. Neural encoding of relative position. *J Exp Psychol Hum Percept Perform*. 37:1032–1050.
- Hoffman DD, Singh M. 1997. Saliency of visual parts. *Cognition*. 63:29–78.
- Hummel JE, Biederman I. 1992. Dynamic binding in a neural network for shape recognition. *Psychol Rev*. 99:480–517.
- Hummel JE, Holyoak KJ. 2003. A symbolic-connectionist theory of relational inference and generalization. *Psychol Rev*. 110:220–264.
- Jolicoeur P. 1985. The time to name disoriented natural objects. *Mem Cognit*. 13:289–303.
- Just MA, Carpenter PA. 1976. Eye fixations and cognitive processes. *Cogn Psychol*. 8:441–480.
- Kamitani Y, Tong F. 2005. Decoding the visual and subjective contents of the human brain. *Nat Neurosci*. 8:679–685.
- Kay KN, Naselaris T, Prenger RJ, Gallant JL. 2008. Identifying natural images from human brain activity. *Nature*. 452:352–355.
- Kayaert G, Biederman I, Op de Beeck HP, Vogels R. 2005. Tuning for shape dimensions in macaque inferior temporal cortex. *Eur J Neurosci*. 22:212–224.
- Kayaert G, Biederman I, Vogels R. 2003. Shape tuning in macaque inferior temporal cortex. *J Neurosci*. 23:3016–3027.
- Kiani R, Esteky H, Mirpour K, Tanaka K. 2007. Object category structure in response patterns of neuronal population in monkey inferior temporal cortex. *J Neurophysiol*. 97:4296–4309.
- Kim JG, Biederman I. 2010. Where do objects become scenes? *Cereb Cortex*. 21:1738–1746.
- Kimia BB. 2003. On the role of medial geometry in human vision. *J Physiol Paris*. 97:155–190.
- Kleiner M, Brainard DH, Pelli DG. 2007. What's new in Psychtoolbox. *Perception*. 36. ECVF Abstract Supplement.
- Konen C, Behrmann M, Nishimura M, Kastner S. 2011. The functional neuroanatomy of object agnosia: a case study. *Neuron*. 71(1):49–60.
- Kourtzi Z, Kanwisher N. 2001. Representation of perceived object shape by the human lateral occipital complex. *Science*. 293:1506–1509.
- Kriegeskorte N, Mur M, Ruff DA, Kiani R, Bodurka J, Esteky H, Tanaka K, Bandettini PA. 2008. Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron*. 60:1126–1141.
- Lades JCV, Buhmann J, Lange J, Malsburg C, Wurtz R, Konen W. 1993. Distortion invariant object recognition in the dynamic link architecture. *IEEE Trans Comput*. 42:300–311.
- Lee TS, Mumford D, Romero R, Lamme VA. 1998. The role of the primary visual cortex in higher level vision. *Vision Res*. 38:2429–2454.
- Mahon BZ, Anzellotti S, Schwarzbach J, Zampini M, Caramazza A. 2009. Category-specific organization in the human brain does not require visual experience. *Neuron*. 63:397–405.
- Marr D, Nishihara HK. 1978. Representation and recognition of the spatial organization of three-dimensional shapes. *Proc R Soc Lond B Biol Sci*. 200:269–294.
- Meyer K, Kaplan JT, Essex R, Webber C, Damasio H, Damasio A. 2010. Predicting visual stimuli on the basis of activity in auditory cortices. *Nat Neurosci*. 13:667–668.
- Op de Beeck HP, Baker CI, DiCarlo JJ, Kanwisher NG. 2006. Discrimination training alters object representations in human extrastriate cortex. *J Neurosci*. 26:13025–13036.
- Op de Beeck HP, Torfs K, Wagemans J. 2008. Perceived shape similarity among unfamiliar objects and the organization of the human object vision pathway. *J Neurosci*. 28:10111–10123.
- Ostwald D, Lam JM, Li S, Kourtzi Z. 2008. Neural coding of global form in the human visual cortex. *J Neurophysiol*. 99:2456–2469.
- Pasupathy A, Connor CE. 1999. Responses to contour features in macaque area V4. *J Neurophysiol*. 82:2490–2502.
- Pasupathy A, Connor CE. 2001. Shape representation in area V4: position-specific tuning for boundary conformation. *J Neurophysiol*. 86:2505–2519.
- Pasupathy A, Connor CE. 2002. Population coding of shape in area V4. *Nat Neurosci*. 5:1332–1338.
- Pelli DG. 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis*. 10:437–442.
- Sereno MI. 1998. Brain mapping in animals and humans. *Curr Opin Neurobiol*. 8:188–194.
- Smith AT, Kossilo P, Williams AL. 2010. The confounding effect of response amplitude on MVPA performance measures. *Neuroimage*. 56(2):525–530.
- Tarr MJ, Pinker S. 1989. Mental rotation and orientation-dependence in shape recognition. *Cogn Psychol*. 21:233–282.
- Thesen S, Heid O, Mueller E, Schad LR. 2000. Prospective acquisition correction for head motion with image-based tracking for real-time fMRI. *Magn Reson Med*. 44:457–465.
- Tsunoda K, Yamane Y, Nishizaki M, Tanifuji M. 2001. Complex objects are represented in macaque inferotemporal cortex by the combination of feature columns. *Nat Neurosci*. 4:832–838.
- Tversky B, Hemenway K. 1984. Objects, parts, and categories. *J Exp Psychol Gen*. 113:169–197.
- Ungerleider LG, Mishkin M. 1982. Two cortical visual systems. In: Ingle DJ, Goodale MA, Mansfield RJW, editors. *Analysis of visual behavior*. Cambridge (MA): MIT Press. p. 549–586.
- Vanrie J, Beatse E, Wagemans J, Sunaert S, Van Hecke P. 2002. Mental rotation versus invariant features in object perception from different viewpoints: an fMRI study. *Neuropsychologia*. 40:917–930.
- Wechsler D. 2004. *The Wechsler intelligence scale for children*. 4th ed. London: Pearson Assessment.
- Williams MA, Dang S, Kanwisher NG. 2007. Only some spatial patterns of fMRI response are read out in task performance. *Nat Neurosci*. 10:685–686.
- Wilson KD, Farah MJ. 2006. Distinct patterns of viewpoint-dependent BOLD activity during common-object recognition and mental rotation. *Perception*. 35:1351–1366.
- Xu Y, Chun MM. 2006. Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature*. 440:91–95.
- Xu X, Yue X, Lescroart MD, Biederman I, Kim JG. 2009. Adaptation in the fusiform face area (FFA): image or person? *Vision Res*. 49:2800–2807.
- Yamane Y, Tsunoda K, Matsumoto M, Phillips AN, Tanifuji M. 2006. Representation of the spatial relationship among object parts by neurons in macaque inferotemporal cortex. *J Neurophysiol*. 96:3147–3156.
- Yue X, Biederman I, Mangini MC, von der Malsburg C, Amir O. 2012. Predicting the psychophysical similarity of faces and non-face complex shapes by image-based measures. *Vision Res*. 55:41–46.
- Yue X, Tjan BS, Biederman I. 2006. What makes faces special? *Vision Res*. 46:3802–3811.
- Zeki SM. 1973. Colour coding in rhesus monkey prestriate cortex. *Brain Res*. 53:422–427.