

available at www.sciencedirect.comjournal homepage: www.elsevier.com/locate/cortex

Research report

Spatial coding and invariance in object-selective cortex

Thomas Carlson^{a,b,c,d,*}, Hinze Hogendoorn^{b,c,d}, Hubert Fonteijn^{c,d}
and Frans A.J. Verstraten^{c,d}

^a Department of Psychology, University of Maryland, College Park, MD, USA

^b Department of Psychology, Vision Sciences Laboratory, Harvard University, Cambridge, MA, USA

^c Helmholtz Institute, Experimental Psychology, Universiteit Utrecht, The Netherlands

^d F.C. Donders Centre for Cognitive Neuroimaging, Radboud University Nijmegen, The Netherlands

ARTICLE INFO

Article history:

Received 31 December 2008

Reviewed 27 March 2009

Revised 31 May 2009

Accepted 28 August 2009

Action editor Maurizio Corbetta

Published online 16 September 2009

Keywords:

Object recognition

fMRI

Spatial coding

Object invariance

ABSTRACT

The present study examined the coding of spatial position in object selective cortex. Using functional magnetic resonance imaging (fMRI) and pattern classification analysis, we find that three areas in object selective cortex, the lateral occipital cortex area (LO), the fusiform face area (FFA), and the parahippocampal place area (PPA), robustly code the spatial position of objects. The analysis further revealed several anisotropies (e.g., horizontal/vertical asymmetry) in the representation of visual space in these areas. Finally, we show that the representation of information in these areas permits object category information to be extracted across varying locations in the visual field; a finding that suggests a potential neural solution to accomplishing translation invariance.

© 2009 Elsevier Srl. All rights reserved.

1. Introduction

Introspectively, the process of recognizing a familiar object is trivial for humans. We quickly and effortlessly recognize objects, and this is easily done across a variety of naturally occurring transformations (e.g., position, size and viewpoint). In terms of the brain, these natural variations present a significant problem to the process of recognition. In particular, the same object can have an infinite number of retinal projections, each of which will evoke unique patterns of neural activity in the cortex. For recognition to occur, these patterns of activation need to be linked to a common representation of a particular object.

One of the natural variations the visual system must deal with is changes in stimulus position on the retina. Human object recognition is generally tolerant to these changes; this is referred to as translation invariance. An early prevailing view was that translation invariance is mechanistically realized by large receptive fields in inferotemporal (IT) cortex, which generalize object identity across changes in spatial position. (Logothetis and Sheinberg, 1996; Tanaka, 1996). A key aspect of this proposal is that IT neurons are highly position tolerant in their response to changes in stimulus position. Recent physiological studies, however, have shown that IT neurons are more sensitive to position than was once thought

* Corresponding author. Department of Psychology, 1145A Biology/Psychology Building, University of Maryland, College Park, MD 20740, USA.

E-mail address: tcarlson@psyc.umd.edu (T. Carlson).

0010-9452/\$ – see front matter © 2009 Elsevier Srl. All rights reserved.

doi:10.1016/j.cortex.2009.08.015

(DiCarlo and Maunsell, 2003; Op De Beeck and Vogels, 2000). These findings have been supported by human functional magnetic resonance imaging (fMRI) studies that show that object selective areas are sensitive to position (Grill-Spector et al., 1999; Hemond et al., 2007; Macevoy and Epstein, 2007; McKyton and Zohary, 2007; Schwarzlose et al., 2008).

The aforementioned studies have demonstrated that position information is robustly coded in the ventral visual pathway. These findings call into question the plausibility of the large receptive field solution to achieve translation invariance. Current research therefore must necessarily explore alternative neural solutions to the question “how do we recognize objects across varying spatial locations?” The present study used functional neuroimaging and pattern classification analyses to address this question. Supporting the results of earlier studies, we show that location information is coded in three regions of interest (ROIs) in the ventral visual pathway: lateral occipital cortex area (LO), fusiform face area (FFA), and parahippocampal place area (PPA). Our analysis further revealed that the representation of visual space in LO, FFA and PPA is non-isotropic. Importantly, we also found that the representation of space and object category information is independent in these areas. Translation invariant object category information can therefore be recovered by integrating object category information across space. Subsequent analyses confirm that a linear classifier can reliably decode object category information, while being tolerant to changes in spatial position (i.e., translation invariance).

2. Materials and methods

2.1. Participants

Nine healthy volunteers (4 female, 5 male) with normal or corrected to normal vision participated in the experiments. Informed written consent was obtained from all volunteers prior to participation. This study was approved by the local ethics committee (Commissie Mensgebonden Onderzoek region Arnhem-Nijmegen, Netherlands).

2.2. Methods and procedure

2.2.1. Visual stimulation

Stimuli were displayed using functions from the Psychtoolbox (Brainard, 1997; Pelli, 1997) in MATLAB (Natick, MA) on an IBM notebook computer. Stimuli were shown on a translucent screen, which subjects viewed through a mirror mounted on the headcoil. Stimuli consisted of 48 grayscale images of faces, scenes, animals, cars, and scrambled images (generated from the images of the objects). Each observer participated in a minimum of 5 functional runs. The displayed images were 8 degrees of visual angle offset from a central fixation point by 7 degrees into the four quadrants of the visual field. Object stimuli (faces, scenes, animals, and cars) were shown in 12 sec blocks followed by 12 sec of scrambled images displayed in the same location (see Fig. 1). Within a block, randomized sequences of images from the same object category (e.g., faces) were shown for 250 msec alternating with 250 msec blank screen with the fixation point. Blocks were displayed to

the four quadrants of the visual field with equal frequency (4 object categories \times 4 locations \times 2 repetitions) in each run.

2.2.2. Experimental task

Subjects fixated a central dot and performed a one-back task responding to repeated presentations of the same image. Repetitions occurred on average twice per block (actual value 1.98). The mean performance across all subjects was 84% correct, with a standard deviation of 4%.

2.2.3. Eye movements

Eye movements were not recorded during the scanning session due to technical limitations. This could potentially be problematic because subjects could have ignored the instructions and fixated the images. However, this can be ruled out by the data. If observers had fixated the images, there would be no variation in retinal location; the images would always be presented to the fovea. It therefore would not be possible to recover retinal location from the ROIs. This is not the case. Retinal location could be decoded from all four ROIs, with near perfect classification performance in early visual areas.

2.2.4. fMRI data acquisition

Data acquisition was performed on a Siemens 3 T Trio scanner at the F.C. Donders Centre for Cognitive Neuroimaging (Nijmegen, The Netherlands). Functional imaging was conducted using an echo-planar imaging (EPI) sequence (TR = 2000, TE = 30msec, flip angle 90°, FOV 192 mm, matrix size 64 \times 64). Scanned volumes were 29 oblique slices (2.5 \times 2.5 mm in plane resolution, 3.5 mm thick; inter-slice gap .35 mm) covering occipital, temporal, and parietal cortex. Functional runs consisted of 388 functional acquisitions (776 sec). The first 4 images of each functional run were discarded to allow for T1 equalization. High resolution T1 weighted anatomical scans were also acquired for each subject.

2.2.5. Data preprocessing

Motion correction and coregistration of functional runs was done in SPM5. The imaging data were then analyzed using Brainvoyager QX version 1.7 (Brain Innovation, Maastricht, The Netherlands). Preprocessing of functional imaging runs consisted of slice scan time correction and linear trend removal.

2.2.6. ROI definitions

ROIs were defined using one, randomly selected functional imaging run. This run was excluded from the classification analysis. LO, FFA, and PPA were localized using the General Linear Model with the following contrasts: LO (cars > scrambled), FFA (faces > cars), and PPA (scenes > cars). LO was defined as a statistically significant ($p < .0001$, uncorrected) cluster of voxels in dorsal lateral occipital cortex [mean volume Right Hemisphere (RH) 2382 mm³, Left Hemisphere (LH) 2268 mm³]. Note, the region referred to as LO in the present study does not include voxels in the posterior fusiform, which together are referred as lateral occipital complex (LOC). FFA and PPA were defined using the same threshold criteria: FFA (mean volume RH 904 mm³, LH 787 mm³); PPA (RH mean volume 1923 mm³, LH 2234 mm³). Examples of the ROIs are shown in

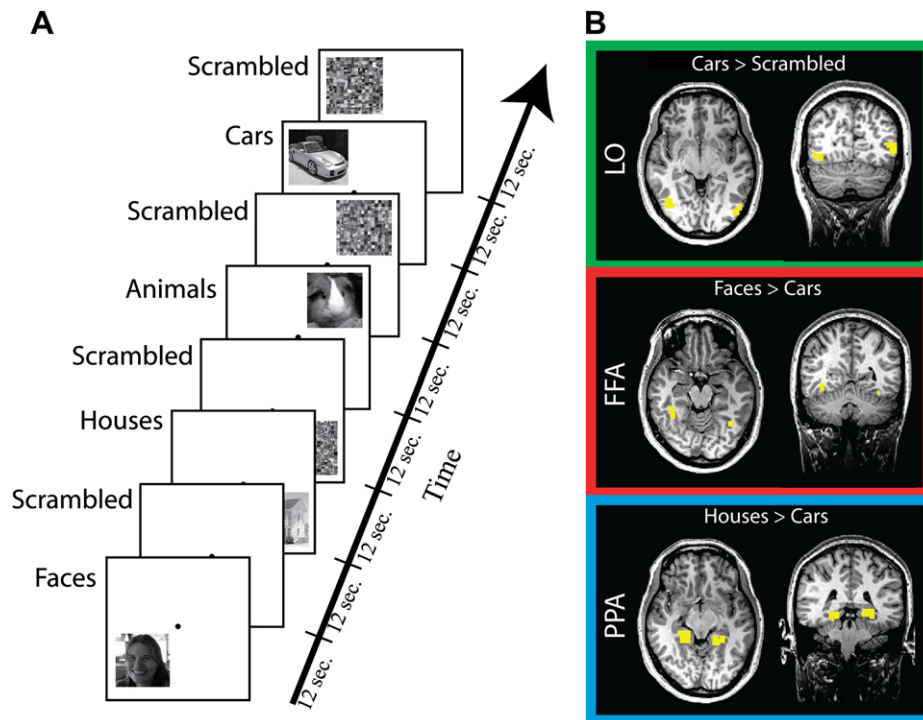


Fig. 1 – A) Experimental paradigm. Four categories of objects (faces, houses, animals, and cars) were displayed with equal frequency to the four quadrants of the visual field. Objects were shown in 12 sec blocks alternating with 12 sec blocks of scrambled images. B) Voxels selected for classification analyses. LO, FFA and PPA shown in a representative subject.

Fig. 1B from a representative subject. The ROI labeled VIS were non-retinotopically defined regions near the calcarine sulcus identified by contrasting stimulated locations in the visual field with non-stimulated locations (e.g., upper left quadrant > lower left, lower right, and the upper right quadrant).

2.2.7. Classification analysis

The time courses from each ROI were extracted and analyzed in MATLAB. The data were deconvolved using a filter given by a gamma function with parameters ($n = 2$, $\tau = 1.25$) in agreement with empirical measures (Boynton et al., 1996). The deconvolution was performed to remove the lag and blurring associated with hemodynamic response to obtain a more accurate representation of neural activity at a given fMRI acquisition. The result of this process is a set of EPI acquisitions labeled by location and category. Principal component analysis (PCA) was performed on the deconvolved data to reduce the dimensionality of data set. This method has effectively been used in a previous study to make single time point predictions with fMRI data using linear discriminant analysis (Carlson et al., 2003). The robustness of this method is further supported in the current dataset in showing near perfect classification performance for decoding object location in early visual areas (see Fig. 2).

2.2.7.1. CATEGORY AND POSITION CLASSIFICATION ANALYSES. Classification was performed using multiple linear discriminant analysis (Duda et al., 2001), in conjunction with a modified leave-one-out procedure. In the analysis, a single test acquisition was removed from the data set. The classifier was

trained using the subset of the data that excluded the test acquisition and all acquisitions from the same block. The exclusion of time points from within the same block was done to remove any effects of serial correlations from the classification analysis. This procedure was repeated until each time point had been tested following a standard leave-one-out procedure.

2.2.7.2. VISUAL FIELD ASYMMETRIES AND CATEGORICAL INTERACTIONS. Linear classifiers were trained and tested on a series of pairwise comparisons between different locations using the above modified leave-one-out procedure. The classifier's performance was evaluated for the following: (1) upper left versus upper right, (2) upper right versus lower right, (3) lower right versus lower left, and (4) lower left versus upper left. In the analysis of categorical interactions and spatial coding location, separate classifiers were trained for each object category.

2.2.7.3. LOCATION SPECIFIC (LS) AND LOCATION INVARIANT (LI) CLASSIFICATION. The LS and LI classification analysis was the same as the initial object category classification analysis except the classifier made a two alternative classification. A single object category was selected (e.g., faces), and the remaining time points were grouped into a single alternative category (i.e., non-faces). For the LS classification, data from each of the four locations was used to train and test four separate classifiers. The LI analysis was done with a single classifier using the data from all four locations, similar to the initial category and position classification analysis (see Section 2.2.7.1).

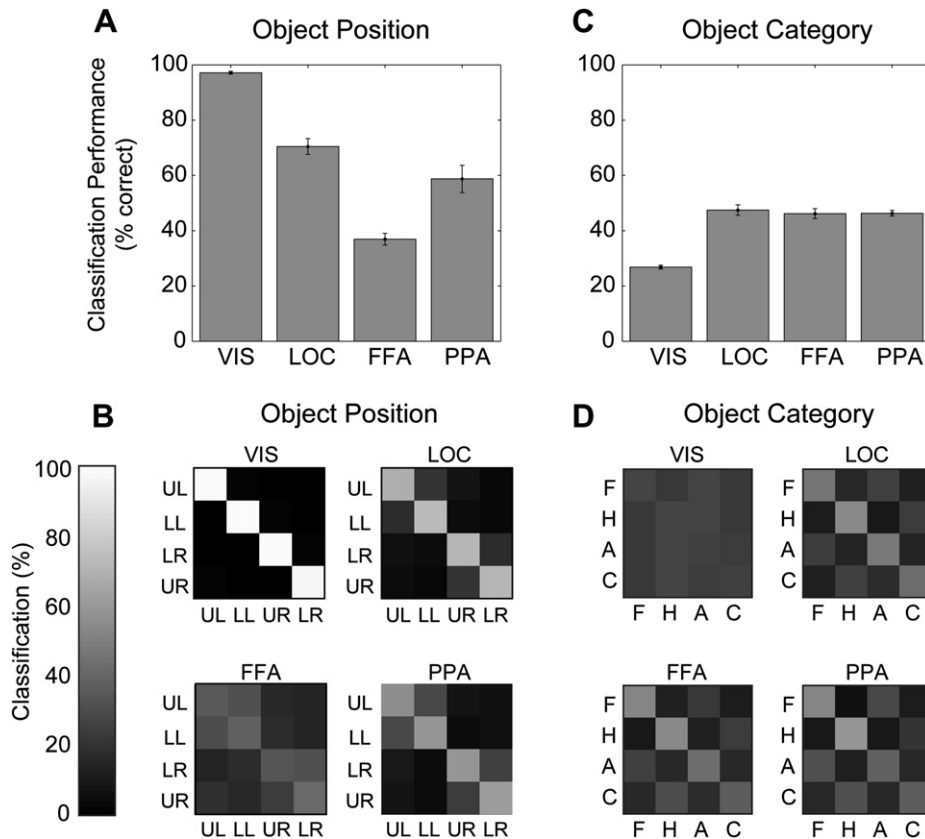


Fig. 2 – Classifications results. **A)** Results from four-way classification analysis decoding object position. Bar graphs show decoding accuracy (chance = 25% correct). Error bars are the Standard Error of the Mean across subjects. **B)** CMs for object position classification analysis. Columns are the location of the object shown to the subject (UL = Upper left quadrant; LR = Lower right quadrant; UR = Upper right quadrant; LL = Lower left quadrant). Rows are the responses from the classifier. The percentages of responses are displayed as grayscale values in cells of the CM. **C)** Results from four-way classification analysis of object category. Bar graphs show decoding accuracy (chance = 25% correct). Error bars are the Standard Error of the Mean across subjects. **D)** CMs for object category classification analysis. Columns are the category of object shown to the subject (F = face; H = house; A = animal; C = car). Rows are the categorical responses from the classifier.

3. Results

3.1. Category and position coding in object-selective cortex

Observers were shown four categories of objects (faces, houses, animals, and cars) displayed to the four quadrants of the visual field in a standard block paradigm (Fig. 1A). Prior to the classification analysis, LO, FFA, and PPA were identified using standard contrasts. Fig. 1B shows the three ROIs in representative subject. The initial analysis also included a fourth ROI (VIS) located in early retinotopic visual areas (~V1).

We first tested if, and if so, to what degree, each of the cortical areas codes the position of objects in the visual field. A linear classifier was trained to decode the quadrant of the visual field in which an object was presented to the observer (upper left, upper right, lower left and lower right; chance performance = 25% correct). Results of the classification analysis are shown in Fig. 2A. Classification performance for object position was near perfect (97% correct) in VIS. The

accurate decoding of spatial position in early visual an area is expected given these areas have, on average, small receptive fields. Above chance performance for recovering the position of the object was also possible in the three object ROIs. The highest performance was LO (69%) followed by PPA (57%), and then FFA (37%). In all three areas, performance was markedly lower than in early visual areas. Still, the ability to recover the location of the object shows that LO, FFA, and PPA have at least a coarse representation of visual space. The confusion matrices (CMs) shown in Fig. 2B provide a more detailed view of the errors in the classification. As can be seen in the figure, the errors were not entirely random in LO, FFA and PPA. The regularities in the CMs show that errors were more likely to occur between locations within the same hemifield (e.g., upper right vs lower right). These regularities are explored further in a subsequent analysis that closer examines anisotropies in the coding of visual space.

We next tested if we could decode the category of object presented to the observer. A linear classifier was trained to recover the category of object (face, house, animal, or car; chance performance = 25% correct) that was presented to the

observer. The results are shown in Fig. 2C. Classification performance was near chance (27%) in early visual areas, thus demonstrating this region contains little explicit information about object category. In contrast, performance was well above chance in LO (47%), FFA (46%), and PPA (46%). This is consistent with a large body of literature showing object oriented processing in these areas. The CMs for the object category classification analysis are shown in Fig. 2D. Again, the resulting CMs showed notable regularities in LO, FFA, and PPA. Here, errors occurred most often between faces and animals, and between cars and houses. The inability to decode object category in early visual areas shows these errors cannot simply be explained by low level image properties. The errors therefore likely reflect commonalities in the manner in which these areas represent visual information, which is somewhat unexpected given the differences in their proposed functional specialization (Kanwisher et al., 1997; Epstein et al., 1999). A second notable point is that location information was not included in the training of the object category classifier. Consequently, the classifier could not simply learn to decode a specific activation pattern (e.g., face in the upper left quadrant). Instead, the classifier was forced to learn a decision criterion for decoding object category that was position invariant. A subsequent analysis examines this “position invariant” decoding of object category further.

3.2. Asymmetries in the representation of space

In early visual areas, the fovea is overrepresented relative to the periphery due to cortical magnification. A non-isotropic representation of visual space with respect to eccentricity has also been observed in object-selective areas (Hasson et al., 2002; Levy et al., 2001, 2004). In addition, several recent studies have documented upper and lower visual field asymmetries in ventral-temporal cortex (Niemeier et al., 2005; Sayres and Grill-Spector, 2008; Schwarzlose et al., 2008). The stimuli in the present study were displayed at the same eccentricity (7 degrees of visual angle) to the four quadrants of the visual field. It therefore possible to test for two types of anisotropies in the coding of visual space: 1) upper or lower visual field asymmetries, and 2) asymmetries along the vertical or horizontal axes.

To test for anisotropies, a new set of classifiers were trained to decode the location of the object presented to the observer for several pairs of locations. Four location pairs were tested: [upper left/lower left], [upper right/lower right], [lower left/lower right], and [upper left/upper right]. Spatial anisotropies were accessed by comparing the classifier’s performance across relevant location pairs. The principal assumption underlying these comparisons is that there will be less overlap in the neuronal representations of two locations in an overrepresented region of a cortical map. Consequently, the evoked activity from the two locations would be more distinct, which would result in better classification performance. The upper/lower asymmetry was tested by comparing performance between the [upper left/upper right] and the [lower left/lower right] location pairs (Fig. 3A). A repeated measures ANOVA [$F(1,7)$] showed a significant difference in LO, with greater accuracy in the lower visual field. The vertical/horizontal asymmetry was tested by comparing the

performance for the upper left/lower left and upper right/lower right pairs to performance for the upper left/upper right and lower left/lower right pairs. A repeated measures ANOVA [$F(1,7)$] found a significant difference in all three areas (Fig. 3A), with better performance for stimuli shown in different hemifields.

Fig. 3B shows projections of the discriminant weights for the four pairs in a representative subject. The discriminant weights are set by the classifier to maximize the separability between the pair (e.g., upper left vs upper right), and therefore can be informative about the distribution of information within an ROI. For example, the weights are positive in one hemisphere and negative in the other hemisphere for the between hemifield classification (1st and 2nd column in figure Fig. 3B). Clearly, this weighting pattern reflects the previously observed contralateral bias in these ROIs (Macevoy and Epstein, 2007; Hemond et al., 2007). This pattern was visually identifiable, to varying degrees, in all of the subjects indicating the contralateral bias is a critical feature for the between hemifield classification. In contrast, we failed to observe any reliable pattern across subjects for the within hemifield classification (3rd and 4th column in Fig. 3B). The failure to observe reliable patterns across subjects indicates that the within hemifield classification relies on more subtle patterns of activation.

3.3. Categorical interactions in the representation of space

FFA and PPA respond preferentially to images of faces and scenes respectively, which is consistent with their proposed specialization for face and scene processing (Epstein et al., 1999; Kanwisher et al., 1997). This preferred response, hypothetically, could convey information that that would improve localization. Alternatively, the mapping of space in FFA and PPA could be fixed across object categories. These two alternatives were tested by individually training separate classifiers for each object category for the locations pairs described above. A repeated measures ANOVA was then used to test for effects of object category on classification performance for spatial position. The analysis found no effect of object category in LO [$F(3,124) = .17, p = .92$], FFA [$F(3,124) = .89, p = .45$] nor PPA [$F(3,124) = .3, p = .83$]. The results therefore indicate the coding of visual space is consistent across object categories.

3.4. Category-selective position-invariant information

How do humans recognize objects across changes in spatial position on the retina? Our initial analysis showed that position invariant object category information could be decoded from LO, FFA and PPA. In our final analysis, we compared a classifier that was trained with knowledge of object location to a position invariant classifier, similar to the one in the initial analysis. Both classifiers were trained to make category specific decisions (e.g., faces vs non-face), distinguishing one category of objects from the others. The first classifier was LS in the training and subsequently tested with data from the same location. Separate classifiers were trained, one for each quadrant of the visual field. The second classifier was trained

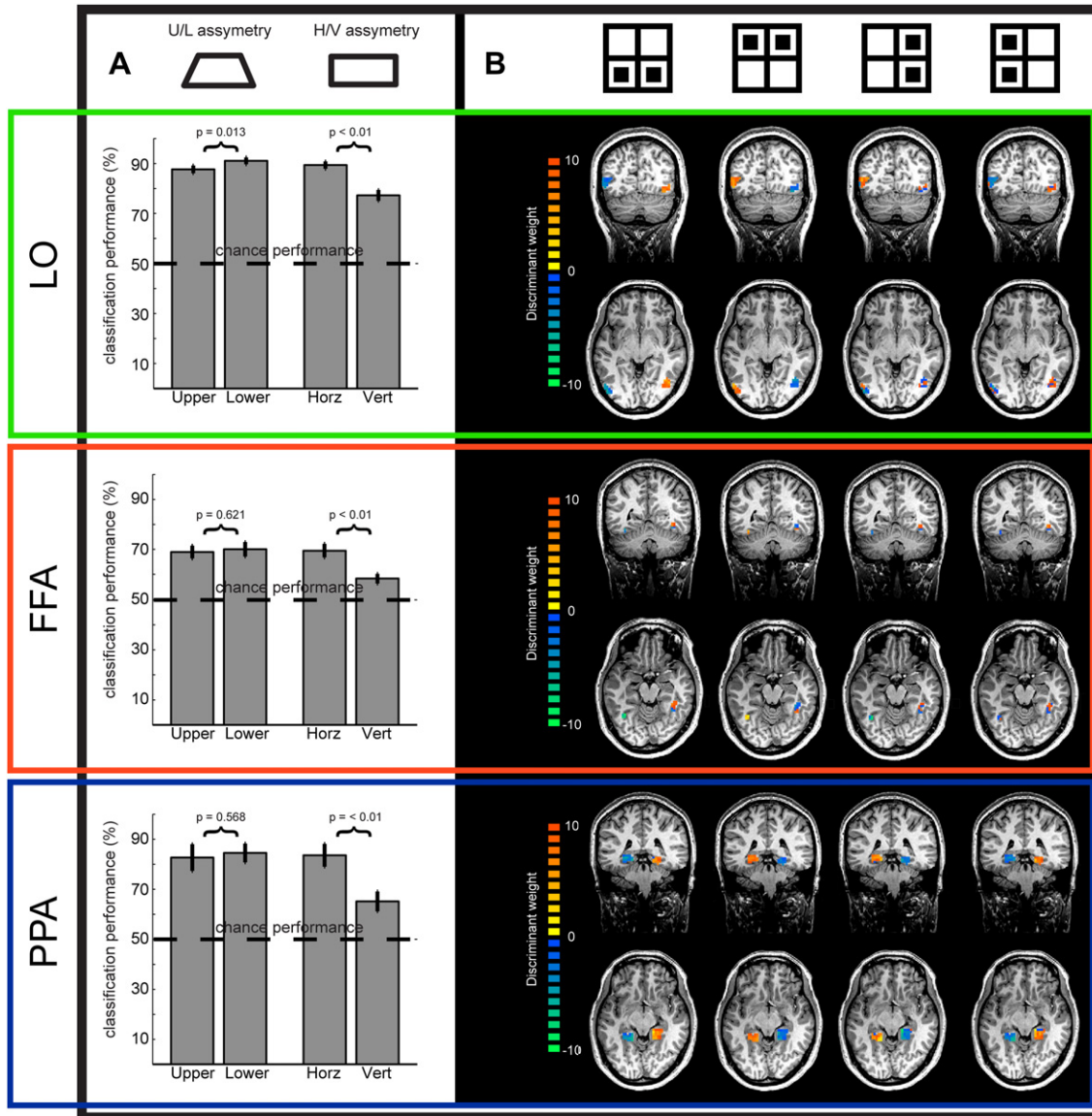


Fig. 3 – Two-way classification analysis decoding object position. A) Bar graphs show the average performance across subjects for the upper/lower (U/L) and vertical/horizontal (V/H) classification. Error bars are the Standard Error of the Mean across subjects. **B)** Projections of the discriminant weights for a representative subject. Weights are arbitrarily scaled between values of -10 and 10 . Columns show the discriminant weights for the four classifiers trained to discriminate the following location pairs: [upper left/lower left], [upper right/lower right], [upper left/upper right], and [lower left/lower right].

to perform the same category specific classification, but was trained with data from all four locations, and also tested with data from all four locations. Importantly, this second classifier needed to learn to decode object category without knowledge of the object's position. We refer to this second classifier as LI.

The results from the analyses are shown in Fig. 4. D-prime is shown for each object category for LS and LI classifiers. Performance, in terms of percent correct, is displayed at the bottom of the bars. The LS and LI classifiers were both above chance for all four categories of objects in all three ROIs. Differences between the classifiers were accessed using a repeated measures ANOVA. There was a significant decrease in performance in LO [$F(1,8) = 14.50$, $p < .01$] and FFA

[$F(1,8) = 10.04$, $p < .01$], with the LS classifier performing slightly better. The difference between the two classifiers was not significant in PPA [$F(1,8) = 1.05$, $p = .34$]. Notably, the difference between the LS and LI classifiers both in terms of d-prime and performance was small. The comparison therefore shows that “LI” category information can be decoded with minimal cost from these areas.

4. Discussion

We found that LO, FFA and PPA code the position of objects in the visual field. This finding was initially demonstrated using

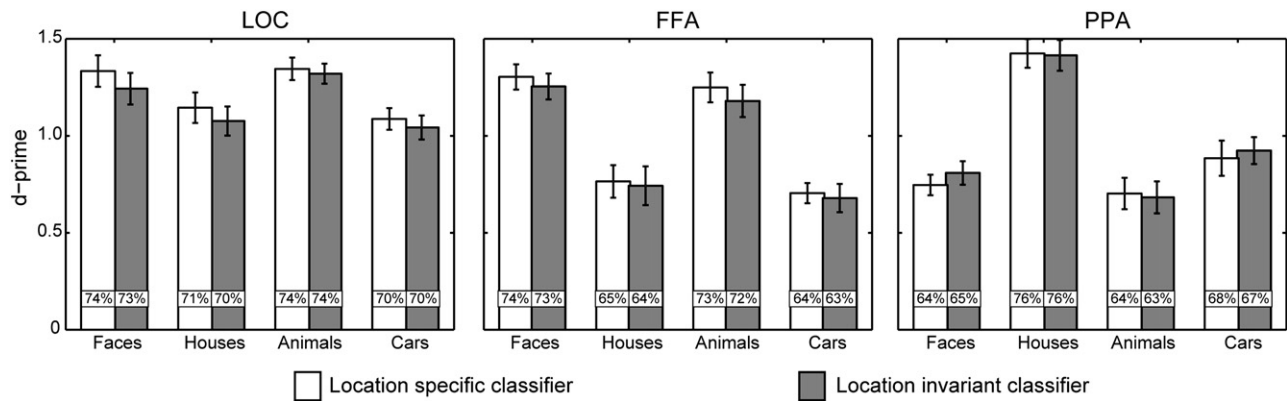


Fig. 4 – Comparison of LS and LI classification. Bar graphs show *d*-prime for the two classifiers for category specific classification (e.g., faces vs non-face). Classification performance is shown at the base of the bars. Error bars are the Standard Error of the Mean across subjects. Although there was a significant effect of LS > LI in two areas, generally performance was high and the difference was small, so there is little to no cost of decoding location invariance.

standard classification methods, and further supported by additional analyses of the representation of space. While our findings are inconsistent with the general notion of ventral-temporal pathway being functionally defined as the so-called “what” pathway (Ungerleider and Mishkin, 1982), they are consistent with a growing body of literature indicating that areas in human ventral-temporal cortex have at least a coarse representation of visual space (Macevoy and Epstein, 2007; Grill-Spector et al., 1999; Larsson and Heeger, 2006; McKyton and Zohary, 2007; Niemeier et al., 2005; Sayres and Grill-Spector, 2008; Schwarzlose et al., 2008).

4.1. The representation of space in object-selective cortex

In LO, FFA and PPA, the classifier performed better for discriminating locations projected to different hemispheres of the brain (i.e., the horizontal/vertical asymmetry). This asymmetry could arise from one, or a combination of factors. First, the asymmetry might be attributed to the nature of the measurement (fMRI). Stimuli presented to the same hemisphere would evoke an activation pattern with larger proportion of the adjacent voxels, and thus separating the two response patterns may be more difficult. This possibility is supported by our projections of the discriminants, in which, the between hemifield classification appeared to principally rely on the global contralateral bias, while the within hemifield classification had to rely on more subtle differences in the activation pattern. A second possibility is the mapping of visual space onto the two hemispheres creates discontinuities in the representation of visual space. While we never actually perceive these gaps, there is behavioral evidence that these anatomical gaps can impact behavior (Carlson et al., 2007; Bar and Biederman, 1999). In support of this possibility, recent behavioral studies examining position coding in object recognition have found hemifield effects that would be predicted by the described horizontal/vertical asymmetry (Kravitz et al., 2008).

The present study and other recent studies have reported a lower visual field bias in LO (Niemeier et al., 2005; Schwarzlose et al., 2008; Sayres and Grill-Spector, 2008). Taken together,

these studies show strong converging evidence for this asymmetry. The absence of this asymmetry in FFA and PPA is of notable interest in that it indicates that this asymmetry is not a ubiquitous property of ventral-temporal cortex. Furthermore, it rules out any explanation based on interactions between cortical areas. For example, primate studies have found an overrepresentation of the lower visual field in parietal areas (Galletti et al., 1999). This asymmetry has been used to explain a lower field advantage in behavioral tasks involving attention (He et al., 1996; Carlson et al., 2007; Rubin et al., 1996). Subjects in our study were instructed to perform a one-back task which presumably requires attention. Based on these observations, one could construct a model in which the lower visual field asymmetry is derived from an interaction between parietal areas and LO. This simple model, however, would fail to explain why FFA and PPA did not also show this asymmetry. In fact, the results of our study indicate researchers should be cautious to attribute the lower visual field advantage to specific early visual areas. To date, this lower visual field bias has only been observed in early visual areas (Liu et al., 2006) and in some areas in the ventral-temporal pathway in humans (Sayres and Grill-Spector, 2008; Schwarzlose et al., 2008; Niemeier et al., 2005). Further research will be required to test if this asymmetry is also present in the dorsal pathway in humans.

Schwarzlose et al. (2008) reported an upper visual field bias in the PPA and a lower visual field bias in FFA. In contrast to their results, we did not find evidence for either of these biases. The principal difference between these two studies was in the assessment of a visual field bias. Schwarzlose et al.’s (2008) evidence was based on the global response of a cortical area. In our study, we accessed these asymmetries based on the ability of the classifier to discriminate locations in the visual field. The lack of agreement between these two reports underscores the importance of validating across analysis techniques.

4.2. Position invariant decoding of object category

Our analysis of visual space found that the representation of space in FFA and PPA did not differ for preferred and

non-preferred object categories. This finding is expected if it is assumed that visual space is represented as a map in these areas. The lack of an interaction between spatial coding and object category does, however, have two interesting corollaries. First, the preferential response to specific categories of objects in FFA and PPA does not improve spatial localization – that is FFA cannot “localize” a face any better than a house or car. Second, this coding scheme bears a resemblance to a recent theoretical model that has been proposed to solve the problem of invariant object recognition (DiCarlo and Cox, 2007). In this model, the joint coding of populations of neurons forms a multidimensional space that represents the many varying properties of an object (e.g., location and pose). Importantly, within this multidimensional space, objects form clusters that are linearly separable from other objects. The use of discriminant analysis and functional neuroimaging data is particularly well suited for testing this model. The response of a single voxel in an fMRI study is an indirect measure of a population of neurons. Pattern classifiers further combine these voxels, which allows for the examination of population responses over large cortical distances. In addition, linear discriminant analysis uses a decision axis to perform the classification similar to that in the model proposed by DiCarlo and Cox (2007). Our analysis showed that indeed a decision axis does exist that can be used to recover object category information that is invariant to location. Importantly, this shows a hypothetical decision neuron could extract categorical information from these areas that is position invariant.

In recent years, there has been an increasing trend to use powerful machine learning classifiers to reveal the information stored in cortical areas (Norman et al., 2006). An important consideration in interpreting the results of these analyses is whether the brain actually makes use of the information. The relevance of this question was recently highlighted in a study that showed that early visual areas could decode two categories of objects, but performance of the classifier did not correlate with human behavior (Williams et al., 2007). The present study found that areas engaged in object oriented processing code spatial position, and further speculates that the problem of invariance might be solved by integrating object information over space. Although speculative, this coding scheme could potentially explain the seminal “what” and “where” pathway findings of Ungerleider and Mishkin (1982), while simultaneously addressing the mounting evidence that “what” pathway neurons robustly code the position of objects. Ungerleider and Mishkin found monkeys were unable to perform the location task after lesioning the so-called “where” pathway. If the intact “what” pathway also codes the location of objects, why then could the monkey not perform the location task? One reason could be that while these areas code the position of objects, this information is inaccessible. This would be true in the coding scheme described above. If neurons were wired to solve the problem of location invariance by integrating object information over space, the output decision neurons would be effectively “blind” to location.

The ability to extract translation invariant categorical information from these areas, however, is only a fractional solution to the problem of invariance. As noted in the

introduction, the visual system must also be able to handle many other sources of image variation (e.g., viewpoint, illumination, etc). Moreover, the visual system must be capable of solving the same problem for exemplars within an object category (e.g., Henry vs Jane). Recent research has shown it is possible to decode exemplars from within categories (Eger et al., 2008; Kriegeskorte et al., 2007). Further study will be required to determine if similar invariant decision axes exist for other sources of variation, and for objects at different levels of description (i.e., category vs identity).

Acknowledgements

This research was supported by a Pionier Grant to F.A.J.V. from the Netherlands Organisation for Scientific Research.

REFERENCES

- Bar M and Biederman I. Localizing the cortical region mediating visual awareness of object identity. *Proceedings of the National Academy of Sciences of the United States of America*, 96: 1790–1793, 1999.
- Boynton GM, Engel SA, Glover GH, and Heeger DJ. Linear systems analysis of functional magnetic resonance imaging in human v1. *The Journal of Neuroscience*, 16: 4207–4221, 1996.
- Brainard DH. The psychophysics toolbox. *Spatial Vision*, 10: 433–436, 1997.
- Carlson TA, Alvarez GA, and Cavanagh P. Quadrantic deficit reveals anatomical constraints on selection. *Proceedings of the National Academy of Sciences of the United States of America*, 104: 13496–13500, 2007.
- Carlson TA, Schrater P, and He S. Patterns of activity in the categorical representations of objects. *Journal of Cognitive Neuroscience*, 15: 704–717, 2003.
- DiCarlo JJ and Cox DD. Untangling invariant object recognition. *Trends in Cognitive Sciences*, 11: 333–341, 2007.
- DiCarlo JJ and Maunsell JH. Anterior inferotemporal neurons of monkeys engaged in object recognition can be highly sensitive to object retinal position. *Journal of Neurophysiology*, 89: 3264–3278, 2003.
- Duda RO, Hart PE, and Stork DG. *Pattern Classification and Scene Analysis*. New York: Wiley, 2001.
- Eger E, Ashburner J, Haynes J-D, Dolan RJ, and Rees G. Functional magnetic resonance imaging activity patterns in human lateral occipital complex carry information about object exemplars within category. *Journal of Cognitive Neuroscience*, 20: 356–370, 2008.
- Epstein R, Harris A, Stanley D, and Kanwisher N. The parahippocampal place area: recognition, navigation, or encoding? *Neuron*, 23: 115–125, 1999.
- Galletti C, Fattori P, Kutz DF, and Gamberini M. Brain location and visual topography of cortical area V6A in the macaque monkey. *European Journal of Neuroscience*, 11: 575–582, 1999.
- Grill-Spector K, Kushnir T, Edelman S, Avidan G, Itzhack Y, and Malach R. Differential processing of objects under various viewing conditions in the human lateral occipital complex. *Neuron*, 24: 187–203, 1999.
- Hasson U, Levy I, Behrmann M, Hendler T, and Malach R. Eccentricity bias as an organizing principle for human high-order object areas. *Neuron*, 34: 479–490, 2002.

- He S, Cavanagh P, and Intriligator J. Attentional resolution and the locus of visual awareness. *Nature*, 383: 334–337, 1996.
- Hemond CC, Kanwisher NG, and Op de Beeck HP. A preference for contralateral stimuli in human object- and face-selective cortex. *PLoS ONE*, 2: e574, 2007.
- Kanwisher N, McDermott J, and Chun MM. The fusiform face area: a module in human extrastriate cortex specialized for face perception. *The Journal of Neuroscience*, 17: 4302–4311, 1997.
- Kravitz DJ, Kriegeskorte N, and Baker CI. Position information throughout the ventral stream. Poster session presented at the annual society for neuroscience meeting. Washington, DC, 2008 November.
- Kriegeskorte N, Formisano E, Sorger B, and Goebel R. Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 104: 20600–20605, 2007.
- Larsson J and Heeger DJ. Two retinotopic visual areas in human lateral occipital cortex. *The Journal of Neuroscience*, 26: 13128–13142, 2006.
- Levy I, Hasson U, Avidan G, Hendler T, and Malach R. Center-periphery organization of human object areas. *Nature Neuroscience*, 4: 533–539, 2001.
- Levy I, Hasson U, Harel M, and Malach R. Functional analysis of the periphery effect in human building related areas. *Human Brain Mapping*, 22: 15–26, 2004.
- Liu T, Heeger DJ, and Carrasco M. Neural correlates of the visual vertical meridian asymmetry. *Journal of Vision*, 6: 1294–1306, 2006.
- Logothetis NK and Sheinberg DL. Visual object recognition. *Annual Review of Neuroscience*, 19: 577–621, 1996.
- Macevoy SP and Epstein RA. Position selectivity in scene- and object-responsive occipitotemporal regions. *Journal of Neurophysiology*, 98: 2089–2098, 2007.
- McKyton A and Zohary E. Beyond retinotopic mapping: The spatial representation of objects in the human lateral occipital complex. *Cerebral Cortex*, 17: 1164–1172, 2007.
- Niemeier M, Goltz HC, Kuchinad A, Tweed DB, and Vilis T. A contralateral preference in the lateral occipital area: Sensory and attentional mechanisms. *Cerebral Cortex*, 15: 325–331, 2005.
- Norman KA, Polyn SM, Detre GJ, and Haxby JV. Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences*, 10: 424–430, 2006.
- Op De Beeck H and Vogels R. Spatial sensitivity of macaque inferior temporal neurons. *The Journal of Comparative Neurology*, 426: 505–518, 2000.
- Pelli DG. The video toolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10: 437–442, 1997.
- Rubin N, Nakayama K, and Shapley R. Enhanced perception of illusory contours in the lower versus upper visual hemifields. *Science*, 271: 651–653, 1996.
- Sayres R and Grill-Spector K. Relating retinotopic and object-selective responses in human lateral occipital cortex. *Journal of Neurophysiology*, 100: 249–267, 2008.
- Schwarzlose RF, Swisher JD, Dang S, and Kanwisher N. The distribution of category and location information across object-selective regions in human visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, 105: 4447–4452, 2008.
- Tanaka K. Inferotemporal cortex and object vision. *Annual Review of Neuroscience*, 19: 109–139, 1996.
- Ungerleider LG and Mishkin M. *Two Visual Pathways*. Cambridge: MIT Press, 1982.
- Williams MA, Dang S, and Kanwisher NG. Only some spatial patterns of fMRI response are read out in task performance. *Nature Neuroscience*, 10: 685–686, 2007.