Scene Vision: Making Sense of What We See

edited by Kestutis Kveraga and Moshe Bar

The MIT Press

From The MIT Press

MITCogNet

MIT Press books may be purchased at special quantity discounts for business or sales promotional use. For information, please email special_sales@mitpress.mit.

This book was set in Times New Roman MT Std 10/13pt by Toppan Best-set Premedia Limited, Hong Kong. Printed and bound in the United States of America.

Library of Congress Cataloging-in-Publication Data.

Scene vision : making sense of what we see / edited by Kestutis Kveraga and Moshe Bar.
pages cm
Includes bibliographical references and index.
ISBN 978-0-262-02785-4 (hardcover : alk. paper) 1. Visual perception. 2. Vision. I. Kveraga, Kestutis. II. Bar, Moshe (Neuroscientist)
BF241.S334 2014
152.14—dc23
2014007216

10 9 8 7 6 5 4 3 2 1

4

Deconstructing Scene Selectivity in Visual Cortex

Reza Rajimehr, Shahin Nasr, and Roger Tootell

In high-order object-processing areas of the ventral visual pathway, discrete clusters of neurons ("modules") respond selectively to specific categories of complex images such as faces (Kanwisher, McDermott, & Chun, 1997; Tsao, Freiwald, Knutsen, Mandeville, & Tootell, 2003; Tsao, Moeller, & Freiwald, 2008), places/scenes (Aguirre, Zarahn, & D'Esposito, 1998; Epstein & Kanwisher, 1998), body parts (Downing, Jiang, Shuman, & Kanwisher, 2001; Grossman & Blake, 2002), and word forms (Cohen et al., 2000). On the other hand, stimuli of a common category often also share low-level visual cues, and correspondingly, it has been reported that many neurons in the inferior temporal (IT) cortex (which is the final stage of the ventral visual pathway) are selective for specific low-level properties, including surface curvature (Janssen, Vogels, Liu, & Orban, 2001; Kayaert, Biederman, & Vogels, 2005), Fourier descriptor shapes (Schwartz, Desimone, Albright, & Gross, 1983), simple geometry (Brincat & Connor, 2004; Kobatake & Tanaka, 1994), nonaccidental features (geons; Vogels, Biederman, Bar, & Lorincz, 2001), diagnostic features (Sigala & Logothetis, 2002), and color (Koida & Komatsu, 2007). Thus, any given category-selective response might be deconstructed into multiple low-level feature selectivities. In fact a recent theory suggests that overlapping continuous maps of simple features give rise to discrete modules that are selective for complex stimuli (Op de Beeck, Haushofer, & Kanwisher, 2008). Selectivity for low-level visual features may be particularly crucial in the processing of scene images. Scenes encompass a virtually infinite range of possible visual stimuli. Selectivity for such a wide range of stimuli may be constructed only by considering some low-level features that are common to images from the scene category. Here we review recent fMRI studies that have reported certain low-level preferences/biases in the scene-responsive areas of visual cortex.

Organization of Scene-Responsive Cortical Areas in Human and Nonhuman Primates

In humans, fMRI studies have described three visual cortical regions that are more active during the presentation of "places" (images of scenes or isolated houses) than during the presentation of other visual stimuli such as faces, objects, body parts, or scrambled scenes (Aguirre et al., 1998; Bar & Aminoff, 2003; Epstein & Kanwisher, 1998; Hasson, Harel, Levy, & Malach, 2003; Maguire, 2001) (figure 4.1, plate 7). These regions have been named for nearby anatomical landmarks as follows: (1) parahippocampal place area (PPA), (2) transverse occipital sulcus (TOS), and (3) retrosplenial cortex (RSC). A recent meta-analysis and comprehensive mapping of scene-related activations suggest that the three scene-responsive regions are actually centered near—but distinct from—the gyri/sulci for which they were originally named (Nasr et al., 2011).

The scene-responsive PPA is typically centered on the lips of the collateral sulcus and adjacent medial fusiform gyrus rather than on the parahippocampal gyrus per se. Although the size of the PPA varies when it is localized with different localizer stimuli, the peak activity is consistently located on the medial fusiform gyrus. More specifically, the fusiform gyrus is subdivided by a shallow sulcus (the middle fusiform sulcus) into a scene-responsive region on the medial fusiform gyrus (PPA) and a face-responsive region on the lateral fusiform gyrus (fusiform face area [FFA]) (Nasr et al., 2011).

The scene-responsive TOS (renamed the occipital place area by Dilks, Julian, Paunov, & Kanwisher, 2013) is typically centered on the nearby lateral occipital gyrus rather than within its namesake, the transverse occipital sulcus. This scene-responsive region lies immediately anterior and ventral to the retinotopically defined area V3A, in/near retinotopic areas V7, V3B, and/or LO-1 (Nasr et al., 2011).

The scene-responsive RSC is a discrete region consistently located in the fundus of the parieto-occipital sulcus, approximately 1 cm from the RSC as defined by Brodmann areas 26, 29, and 30. This scene-responsive region is located immediately adjacent to V1 in what would otherwise be the peripheral representation of dorsal V2 (Nasr et al., 2011).

The functional connectivity of these scene areas has been tested during the resting-state fMRI (Baldassano, Beck, & Fei-Fei, 2013). The RSC and TOS show differentiable functional connections with the anterior-medial and posterior-lateral parts of the PPA, respectively. Each of these areas is also functionally connected with specific parts of the cortex. The RSC shows connections with the superior frontal sulcus (Brodmann areas 8/9) and the peripheral representation of early visual areas V1 and V2. The TOS shows connections with the intraparietal sulcus, the lateral occipital complex, and retinotopic early visual areas. The PPA shows connections with the lateral occipital complex and the peripheral representation of early visual areas V1 and V2.

Corresponding (presumptively homologous) scene-responsive regions have been identified by use of fMRI in awake macaque monkeys (Nasr et al., 2011; Rajimehr, Devaney, Bilenko, Young, & Tootell, 2011) (figure 4.1, plate 7). These studies used identical stimuli and largely overlapping fMRI procedures in human and monkey scans so that a relatively direct comparison between human and monkey maps was possible. Mirroring the arrangement of the human FFA and PPA (which are adjacent to each other in cortex), the presumptive monkey homologue of the human PPA (mPPA) is located adjacent to the most prominent face patch in the IT cortex. This location is immediately anterior to area TEO. The monkey TOS (mTOS) includes the region predicted from the human maps (macaque V4d), extending posteriorly into V3A. A possible monkey homologue of the human RSC lies in the medial bank, near the peripheral V1.

In addition to mPPA, a recent study has reported two other “place patches” in macaque ventral temporal cortex (Kornblith, Cheng, Ohayon, & Tsao, 2013). These patches, the lateral place patch (LPP) and the medial place patch (MPP), are located in the occipitotemporal sulcus and the parahippocampal gyrus, respectively. The LPP contains a large concentration of scene-selective single units, with individual units coding specific scene parts. Based on microstimulation, the LPP is connected with the MPP and with extrastriate visual areas V4v and DP.

Retinotopic Selectivity in Scene-Responsive Areas

Early electrophysiological studies suggested that neurons in the IT cortex have large receptive fields (>20°) (Desimone & Gross, 1979; Richmond, Wurtz, & Sato, 1983). Those studies emphasized that the positional information is lost at progressively higher stages of the ventral visual pathway, and neurons become selective for visual features and objects independent of their locations in the visual field (Ito, Tamura, Fujita, & Tanaka, 1995; Lueschow, Miller, & Desimone, 1994). However, more recent studies of the IT cortex have reported the presence of small receptive fields (<5°) even in the anterior IT cortex (DiCarlo & Maunsell, 2003; Op De Beeck & Vogels, 2000). In fact there is a wide distribution of receptive field sizes in the IT cortex, ranging from 3° to 25° with a mean size of 10° (Op De Beeck & Vogels, 2000). These data are consistent with the idea that representations in the IT cortex are position dependent. This position sensitivity could be considered a low-level selectivity for object-selective IT neurons.

Analogously, early human fMRI studies distinguished between retinotopic and nonretinotopic cortex (e.g., Grill-Spector et al., 1998; Halgren et al., 1999). Those studies described retinotopic maps in occipital visual areas such as V1, V2, V3, V3A/B, hV4, and V5/hMT+ (e.g., Brewer, Liu, Wade, & Wandell, 2005; DeYoe et al., 1996; Engel, Rumelhart, Wandell, & Lee, 1994; Huk, Dougherty, & Heeger, 2002; Sereno et al., 1995; Tootell et al., 1997) but failed to find consistent retinotopy in higher-level areas of the ventral visual pathway—perhaps due to technical limitations. With technical advancements in neuroimaging and better stimulus designs, recent fMRI studies have reported retinotopic maps located beyond (anterior to) V4. Such maps have been identified in object-selective lateral occipital cortex (Larsson & Heeger, 2006; Sayres & Grill-Spector, 2008) and within regions selective for object categories such as body parts (Weiner & Grill-Spector, 2011). Distributed positional information from multivoxel pattern analysis has also been reported in almost all category-selective

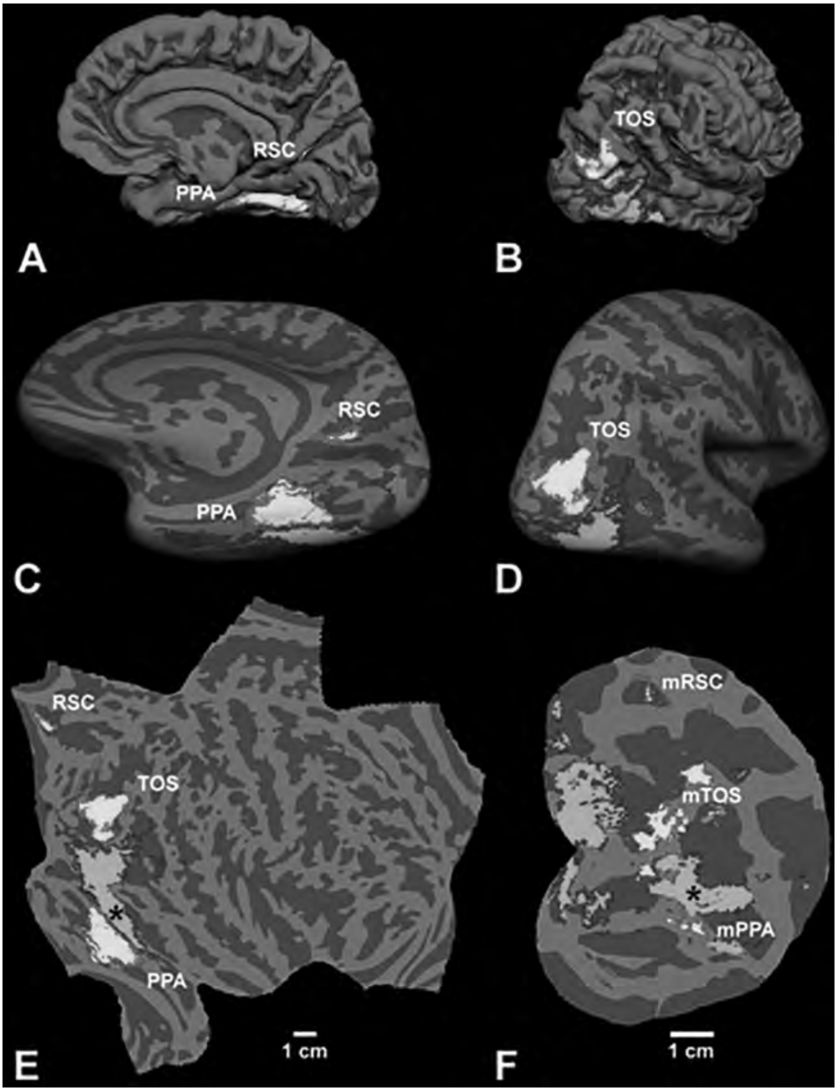

Figure 4.1 (plate 7)
Overall view of scene-responsive areas in human and monkey visual cortex. Both species fixated the center of a screen during block-designed presentation of identical scene versus face-localizing stimuli. In the human data (panels A–E), relatively higher activity to scenes versus faces is shown in red/yellow versus blue/cyan, respectively (minimum = $p < 10^{-10}$; maximum = $p < 10^{-30}$). The human map is a group average of both the functional and the anatomical data ($n = 10$) in cortical surface format. The right hemisphere is illustrated. Panels A and B show the medial and lateral-posterior views of folded cortex, respectively. Panels C and D show corresponding views of inflated cortex, and panel E shows the flattened view. For comparison, panel F shows the flattened activity map from a macaque monkey viewing the same stimuli (minimum = $p < 10^{-5}$; maximum = $p < 10^{-10}$). In both species presumptive corresponding scene-responsive areas are named in white (preceded by “m” in the macaque map). Adapted from Nasr et al. (2011).

areas (Carlson, Hogendoorn, Fonteijn, & Verstraten, 2011; Cichy, Chen, & Haynes, 2011; Kravitz, Kriegeskorte, & Baker, 2010; Schwarzlose, Swisher, Dang, & Kanwisher, 2008). Thus, low-level selectivity for the retinotopic location of visual stimuli is preserved in higher-level areas in human IT cortex.

Using retinotopic mapping combined with an attentional tracking paradigm, Arcaro and colleagues reported two new retinotopic maps anterior to the VO cluster within the posterior parahippocampal cortex (PHC), referred to as PHC-1 and PHC-2 (Arcaro, McMains, Singer, & Kastner, 2009). Each PHC area contains a complete representation of the contralateral visual field with a bias for stimuli in the upper visual field (see also Schwarzlose et al., 2008). Both areas are heavily overlapped with the functionally defined area PPA (Arcaro et al., 2009); this suggests a position-dependent coding of scenes in the PPA. The scene-responsive TOS also has retinotopic selectivity because it is located within the retinotopic extrastriate cortex in both humans and macaques (Nasr et al., 2011).

Scene-related areas, including the PPA and TOS, also show a strong preference for stimuli presented in the peripheral visual field. In a series of experiments Malach and colleagues reported an association between category selectivity and eccentricity bias in high-order object areas; face areas were associated with central visual-field bias, whereas scene areas were associated with peripheral visual-field bias (Hasson et al., 2003; Levy, Hasson, Avidan, Hendler, & Malach, 2001; Malach, Levy, & Hasson, 2002). Retinotopic eccentricity might be an organizing principle of object representations in these areas, and it can be tightly linked to acuity demands (Hasson, Levy, Behrmann, Hendler, & Malach, 2002). It is conceivable that scene-related processes such as spatial navigation and texture segregation depend crucially on large-scale integration, and thus, these functions might be better served by a strong association with peripheral, low-magnification representations.

A Preference for High Spatial Frequencies in Scene-Responsive Areas

In the lower-tier (occipital) visual cortex (e.g., in V1, V2, and V3), the sensitivity to spatial frequency covaries systematically with the retinotopic representation of visual field eccentricity. That is, the foveal/parafoveal cortex in these areas shows a preference for higher spatial frequencies (Henriksson, Nurminen, Hyvärinen, & Vanni, 2008; Sasaki et al., 2001). Recently, the sensitivity to spatial frequency has been tested in the higher-tier areas of the ventral visual pathway (Awasthi, Sowman, Friedman, & Williams, 2013; Rajimehr et al., 2011; Zeidman, Mullally, Schwarzkopf, & Maguire, 2012). One study (Rajimehr et al., 2011) manipulated the spatial frequency in a variety of stimuli and found a significant preference for high spatial frequencies in the scene-responsive area PPA. The high-spatial-frequency bias in the PPA was demonstrated using high-pass-filtered scene, face, and even checkerboard stimuli (figure 4.2, plate

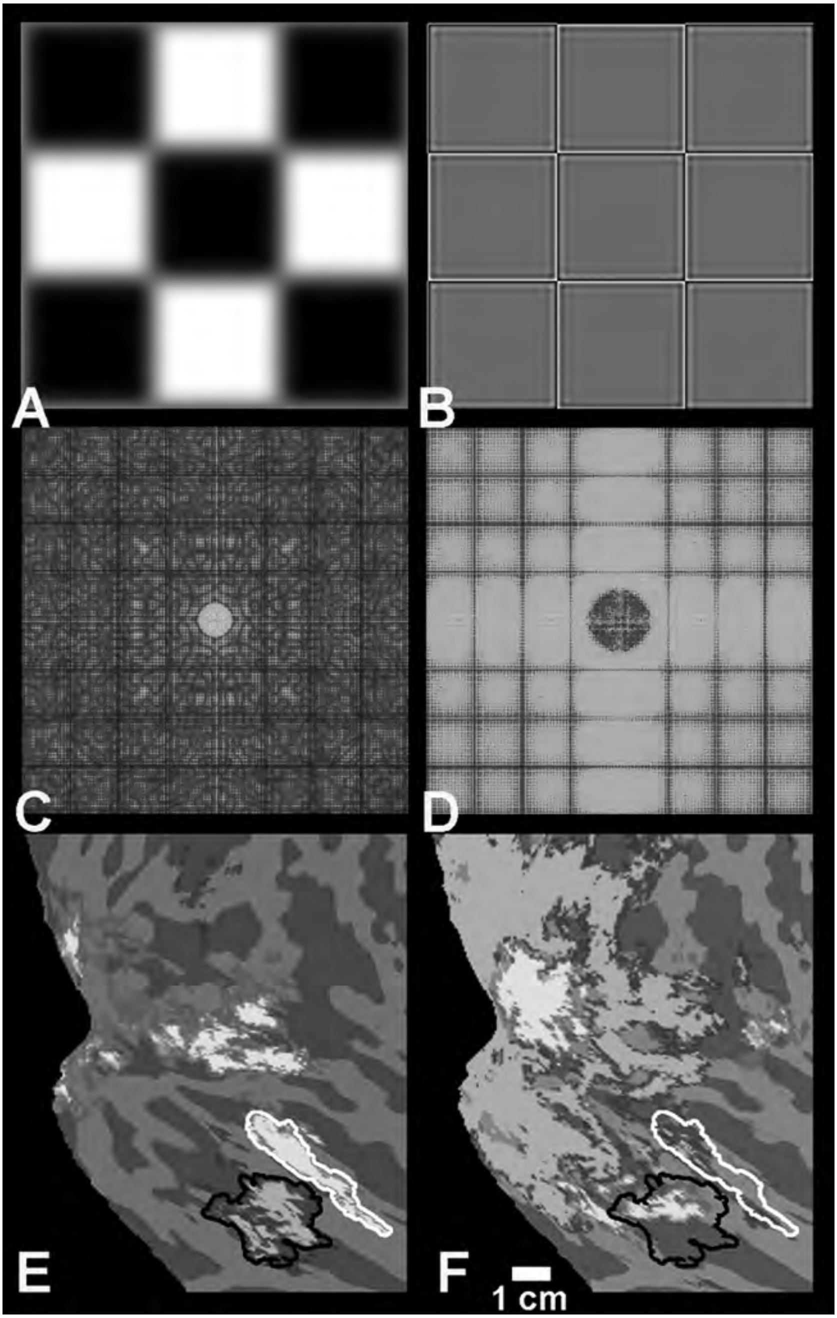

Figure 4.2 (plate 8)
High-pass-filtered checkerboard images selectively activate the PPA. Panels A and B show examples of low-spatial-frequency (low-SF) and high-spatial-frequency (high-SF) checkerboards. Panels C and D show the FFT amplitude spectra of low-SF and high-SF checkerboards. Panel E shows the FFA (indicated by white border) and PPA (indicated by black border), localized based on a comparison between faces and places, in the averaged map of four human subjects. The group-averaged activity map is displayed on a flattened view of the right occipitotemporal cortex. Panel F shows the comparison of activity between high-SF (yellow/red) and low-SF (cyan/blue) checkerboards. This comparison revealed a high-SF bias within the PPA. If anything, the opposite bias was found in parts of the FFA. The maps are significant at the threshold of $p < 10^{-2}$. Adapted from Rajimehr et al. (2011).

8). This bias was more prominent in the posterior-lateral part of the PPA. The PPA also showed a higher response to unfiltered natural scenes that were dominated by high spatial frequencies. This study (Rajimehr et al., 2011) used identical stimuli in monkeys and found that the mPPA (the apparent homologue of PPA in monkeys) also has a preference for high spatial frequencies.

An image analysis suggests that scenes have more spatial discontinuities (in the form of high-spatial-frequency components) compared to other object categories, such as faces (Rajimehr et al., 2011). Thus, a low-level sensitivity to high spatial frequencies in the PPA can be particularly useful for detecting edges, object borders, and scene details during spatial perception and navigation. Furthermore, there may be an evolutionary advantage for the PPA to be preferentially tuned for high spatial frequencies (e.g., in facilitating the detection of food/predators in visually complex environments).

A Cardinal Orientation Bias in Scene-Responsive Areas

Human vision is more sensitive to contours at cardinal (horizontal and vertical) orientations compared to oblique orientations, a phenomenon called the "oblique effect" (Appelle, 1972; Mach, 1861). Because the oblique effect is linked to stimulus orientation, and orientation-selective cells are common in V1, prior psychophysical and physiological experiments have often hypothesized a neural correlate of the perceptual oblique effect in lower-level retinotopic visual cortex (Vogels & Orban, 1985). However, fMRI studies in humans have not reported a consistent activity bias for cardinal orientations in V1 (Freeman, Brouwer, Heeger, & Merriam, 2011; but see Furmanski & Engel, 2000; Swisher et al., 2010).

A recent study suggests that the oblique effect may be related to scene processing (Nasr & Tootell, 2012). The link between the oblique effect and scene processing is supported by ecological evidence. Image statistics confirm that many scenes are dominated by cardinal orientations (Torralba & Oliva, 2003). Such a statistical bias is present not only in carpentered environments (such as cityscapes, buildings, and indoor scenes) but also in some natural scenes, often due to the orthogonal influences of gravity and/or phototropism. Consistent with this idea, the scene-responsive area PPA shows a stronger fMRI response to cardinal (compared to oblique) orientations (Nasr & Tootell, 2012) (figure 4.3, plate 9). This low-level orientation bias in the PPA can be observed even for simple geometrical stimuli such as arrays of overlapping squares or arrays of line segments (Nasr & Tootell, 2012).

Conclusion

Here we reviewed evidence for selectivity to low-level visual features in the scene-responsive areas, particularly in the PPA. It is important to define such features for

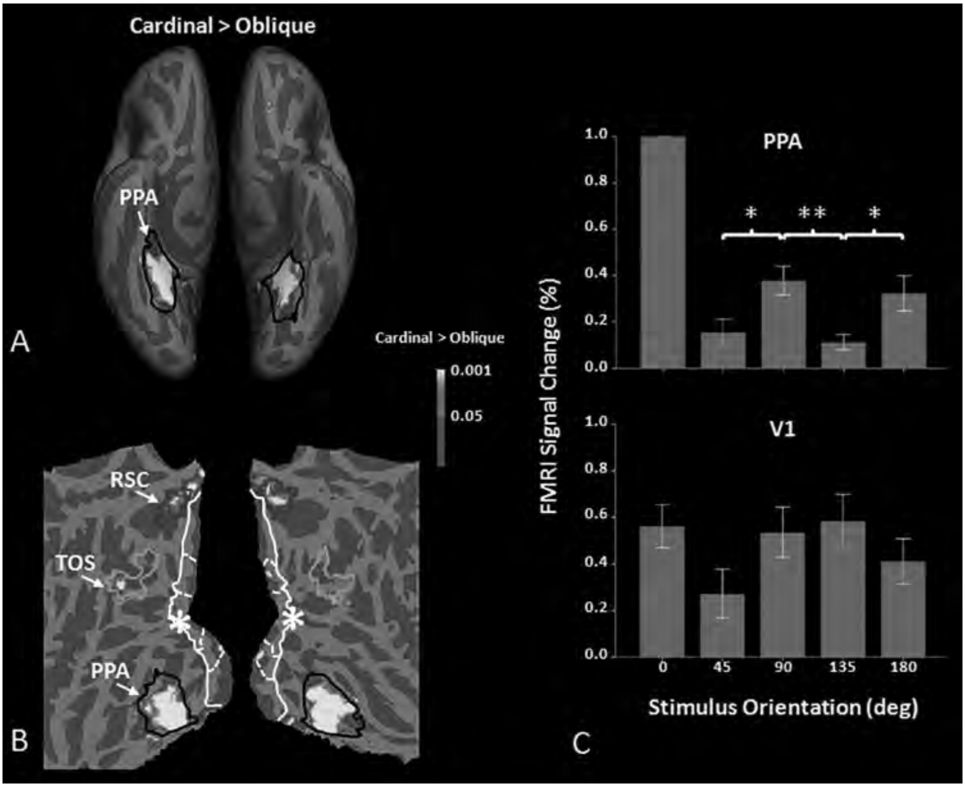

Figure 4.3 (plate 9)
Scenes with dominant power at cardinal orientations selectively activate the PPA. Panel A shows group-averaged fMRI activity (*n* = 13; random-effects analysis) for spatially filtered scenes with dominant power at cardinal versus oblique orientations. The map shows a ventral view of the inflated (panel A) and flattened (panel B) cortical surfaces. The borders of the PPA, TOS, RSC, and V1 are shown using black, green, blue, and white lines, respectively. White dashed lines indicate the peripheral visual field representation, and white asterisks indicate the foveal representation. Panel C shows the region-of-interest analysis in the PPA and V1. Orientation differences were significant only in the PPA (*$p < 0.05$; **$p < 0.01$). Adapted from Nasr et al. (2012).

both practical and conceptual reasons. At a practical level it is important to define such variables to avoid uncontrolled stimulus influences on otherwise carefully controlled tests of higher-order variables. More conceptually, such lower-level variables may well serve as intermediate "building blocks" for higher-order selectivity and thus clarify the nature of further higher-order variables. The latter is particularly important because after its initial characterization as a scene-selective area, the PPA has been reinterpreted as selective for a wide range of other higher-order categories, including tools (Chao, Haxby, & Martin, 1999), single houses (Tootell et al., 2008), inanimate objects (Kriegeskorte et al., 2008), big objects (Konkle & Oliva, 2012; Troiani, Stigliani, Smith, & Epstein, 2014), object ensembles, and surface textures (Cant & Xu, 2012). Due to this wide range of stimulus selectivities in the PPA, we have used the term "scene-responsive" (rather than "scene-selective") when referring to the PPA and other scene areas.

Evidence to date demonstrates that multiple lower-level variables influence the fMRI response properties in the PPA. Retinotopically, studies have reported that this area shows a bias for stimuli in the peripheral (as opposed to foveal) visual field (e.g., Levy et al., 2001) and a bias for upper visual field (Arcaro et al., 2009). It can be argued that to the extent that a response is specific to portions of the visual field, it cannot be strictly selective for a given visual category per se.

More globally, it has been shown that the PPA is also selective for high spatial frequencies and cardinal orientations. It is possible that PPA and other scene-responsive areas also show selectivity to other low-level features such as specific line junctions or specific shapes. For instance, the fact that the PPA responds more strongly to cubes than to spheres (Rajimehr et al., 2011) may be partly due to the presence of right angles in the "cube" stimuli. Future studies would be needed to test this and other low-level feature selectivities in these areas. Demonstration of low-level selectivity in the scene areas makes the single-cell studies of scene processing more tractable, as these low-level features can be parametrically manipulated in well-defined stimulus spaces.

Acknowledgments

This research was supported by National Institutes of Health (NIH) Grants R01 MH67529 and R01 EY017081 to R.B.H.T., the Martinos Center for Biomedical Imaging, the NCRR, and the MIND Institute.

References

Aguirre, G. K., Zarahn, E., & D'Esposito, M. (1998). An area within human ventral cortex sensitive to "building" stimuli: Evidence and implications. *Neuron*, *21*, 373–383.

Appelle, S. (1972). Perception and discrimination as a function of stimulus orientation: The oblique effect in man and animals. *Psychological Bulletin*, *78*, 266–278.

Arcaro, M. J., McMains, S. A., Singer, B. D., & Kastner, S. (2009). Retinotopic organization of human ventral visual cortex. *Journal of Neuroscience*, *29*(34), 10638–10652.

Awasthi, B., Sowman, P. F., Friedman, J., & Williams, M. A. (2013). Distinct spatial scale sensitivities for early categorization of faces and places: Neuromagnetic and behavioral findings. *Frontiers in Human Neuroscience*, *7*(91).

Baldassano, C., Beck, D. M., & Fei-Fei, L. (2013). Differential connectivity within the parahippocampal place area. *NeuroImage*, *75*, 228–237.

Bar, M., & Aminoff, E. M. (2003). Cortical analysis of visual context. *Neuron*, *38*, 347–358.

Brewer, A. A., Liu, J., Wade, A. R., & Wandell, B. A. (2005). Visual field maps and stimulus selectivity in human ventral occipital cortex. *Nature Neuroscience*, *8*(8), 1102–1109.

Brincat, S. L., & Connor, C. E. (2004). Underlying principles of visual shape selectivity in posterior inferotemporal cortex. *Nature Neuroscience*, *7*(8), 880–886.

Cant, J. S., & Xu, Y. (2012). Object ensemble processing in human anterior-medial ventral visual cortex. *Journal of Neuroscience*, *32*(22), 7685–7700.

Carlson, T., Hogendoorn, H., Fonteijn, H., & Verstraten, F. A. (2011). Spatial coding and invariance in object-selective cortex. *Cortex*, *47*(1), 14–22.

Chao, L. L., Haxby, J. V., & Martin, A. (1999). Attribute-based neural substrates in temporal cortex for perceiving and knowing about objects. *Nature Neuroscience*, *2*(10), 913–919.

Cichy, R. M., Chen, Y., & Haynes, J. D. (2011). Encoding the identity and location of objects in human LOC. *NeuroImage*, *54*(3), 2297–2307.

Cohen, L., Dehaene, S., Naccache, L., Lehéricy, S., Dehaene-Lambertz, G., Hénaff, M. A., et al. (2000). The visual word form area: Spatial and temporal characterization of an initial stage of reading in normal subjects and posterior split-brain patients. *Brain*, *123*(2), 291–307.

Desimone, R., & Gross, C. G. (1979). Visual areas in the temporal cortex of the macaque. *Brain Research*, *178*(2), 363–380.

DeYoe, E. A., Carman, G. J., Bandettini, P., Glickman, S., Wieser, J., Cox, R., et al. (1996). Mapping striate and extrastriate visual areas in human cerebral cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *93*(2), 2382–2386.

DiCarlo, J. J., & Maunsell, J. H. (2003). Anterior inferotemporal neurons of monkeys engaged in object recognition can be highly sensitive to object retinal position. *Journal of Neurophysiology*, *89*(6), 3264–3278.

Dilks, D. D., Julian, J. B., Paunov, A. M., & Kanwisher, N. (2013). The occipital place area (OPA) is causally and selectively involved in scene perception. *Journal of Neuroscience*, *33*(4), 1331–1336.

Downing, P. E., Jiang, Y., Shuman, M., & Kanwisher, N. (2001). A cortical area selective for visual processing of the human body. *Science*, *293*(5539), 2470–2473.

Engel, S. A., Rumelhart, D. E., Wandell, B. A., & Lee, A. T. (1994). fMRI of human visual cortex. *Nature*, *369*, 525.

Epstein, R. A., & Kanwisher, N. (1998). A cortical representation of the local visual environment. *Nature*, *392*(6676), 598–601.

Freeman, J., Brouwer, G. J., Heeger, D. J., & Merriam, E. P. (2011). Orientation decoding depends on maps, not columns. *Journal of Neuroscience*, *31*(13), 4792–4804.

Furmanski, C. S., & Engel, S. A. (2000). An oblique effect in human primary visual cortex. *Nature Neuroscience*, *3*(6), 535–536.

Grill-Spector, K., Kushnir, T., Hendler, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Human Brain Mapping*, *6*(4), 316–328.

Grossman, E. D., & Blake, R. (2002). Brain areas active during visual perception of biological motion. *Neuron*, *35*(6), 1167–1175.

Halgren, E., Dale, A. M., Sereno, M. I., Tootell, R. B. H., Marinkovic, K., & Rosen, B. R. (1999). Location of human face-selective cortex with respect to retinotopic areas. *Human Brain Mapping*, *7*(1), 29–37.

Hasson, U., Harel, M., Levy, I., & Malach, R. (2003). Large-scale mirror-symmetry organization of human occipito-temporal object areas. *Neuron*, *37*, 1027–1041.

Hasson, U., Levy, I., Behrmann, M., Hendler, T., & Malach, R. (2002). Eccentricity bias as an organizing principle for human high-order object areas. *Neuron*, *34*(3), 479–490.

Henriksson, L., Nurminen, L., Hyvärinen, A., & Vanni, S. (2008). Spatial frequency tuning in human retinotopic visual areas. *Journal of Vision, 8*(10), 1–13.

Huk, A. C., Dougherty, R. F., & Heeger, D. J. (2002). Retinotopy and functional subdivision of human areas MT and MST. *Journal of Neuroscience*, *22*(16), 7195–7205.

Ito, M., Tamura, H., Fujita, I., & Tanaka, K. (1995). Size and position invariance of neuronal responses in monkey inferotemporal cortex. *Journal of Neurophysiology*, *73*(1), 218–226.

Janssen, P., Vogels, R., Liu, Y., & Orban, G. A. (2001). Macaque inferior temporal neurons are selective for three-dimensional boundaries and surfaces. *Journal of Neuroscience*, *21*(23), 9419–9429.

Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, *17*(11), 4302–4311.

Kayaert, G., Biederman, I., & Vogels, R. (2005). Representation of regular and irregular shapes in macaque inferotemporal cortex. *Cerebral Cortex*, *15*(9), 1308–1321.

Kobatake, E., & Tanaka, K. (1994). Neuronal selectivities to complex object features in the ventral visual pathway of the macaque cerebral cortex. *Journal of Neurophysiology*, *71*(3), 856–867.

Koida, K., & Komatsu, H. (2007). Effects of task demands on the responses of color-selective neurons in the inferior temporal cortex. *Nature Neuroscience*, *10*, 108–116.

Konkle, T., & Oliva, A. (2012). A real-world size organization of object responses in occipito-temporal cortex. *Neuron*, *74*(6), 1114–1124.

Kornblith, S., Cheng, X., Ohayon, S., & Tsao, D. Y. (2013). A network for scene processing in the macaque temporal lobe. *Neuron*, *79*(4), 766–781.

Kravitz, D. J., Kriegeskorte, N., & Baker, C. I. (2010). High-level visual object representations are constrained by position. *Cerebral Cortex*, *20*(12), 2916–2925.

Kriegeskorte, N., Mur, M., Ruff, D., Kiani, R., Bodurka, J., Esteky, H., et al. (2008). Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron*, *60*(6), 1126–1141.

Larsson, J., & Heeger, D. J. (2006). Two retinotopic visual areas in human lateral occipital cortex. *Journal of Neuroscience*, *26*(51), 13128–13142.

Levy, I., Hasson, U., Avidan, G., Hendler, T., & Malach, R. (2001). Center-periphery organization of human object areas. *Nature Neuroscience*, *4*(5), 533–539.

Lueschow, A., Miller, E. K., & Desimone, R. (1994). Inferior temporal mechanisms for invariant object recognition. *Cerebral Cortex*, *4*(5), 523–531.

Mach, E. (1861). Über das Sehen von Lagen und Winkeln durch die Bewegung des Auges. *Sitzungsberichte der Kaiserlichen Akademie der Wissenschaften*, *43*(2), 215–224.

Maguire, E. A. (2001). The retrosplenial contribution to human navigation: A review of lesion and neuroimaging findings. *Scandinavian Journal of Psychology*, *42*, 225–238.

Malach, R., Levy, I., & Hasson, U. (2002). The topography of high-order human object areas. *Trends in Cognitive Sciences*, *6*(4), 176–184.

Nasr, S., Liu, N., Devaney, K. J., Yue, X., Rajimehr, R., Ungerleider, L. G., et al. (2011). Scene-selective cortical regions in human and non-human primates. *Journal of Neuroscience*, *31*(39), 13771–13785.

Nasr, S., & Tootell, R. B. H. (2012). A cardinal orientation bias in scene-selective visual cortex. *Journal of Neuroscience*, *32*(43), 14921–14926.

Op de Beeck, H., Haushofer, J., & Kanwisher, N. G. (2008). Interpreting fMRI data: Maps, modules and dimensions. *Nature Reviews Neuroscience*, *9*(2), 123–135.

Op De Beeck, H., & Vogels, R. (2000). Spatial sensitivity of macaque inferior temporal neurons. *Journal of Comparative Neurology*, *426*(4), 505–518.

Rajimehr, R., Devaney, K. J., Bilenko, N. Y., Young, J. C., & Tootell, R. B. H. (2011). The parahippocampal place area responds preferentially to high spatial frequencies in humans and monkeys. *PLoS Biology*, *9*(4), e1000608.

Richmond, B. J., Wurtz, R. H., & Sato, T. (1983). Visual responses of inferior temporal neurons in awake rhesus monkey. *Journal of Neurophysiology*, *50*(6), 1415–1432.

Sasaki, Y., Hadjikhani, N., Fischl, B., Liu, A. K., Marret, S., Dale, A. M., et al. (2001). Local and global attention are mapped retinotopically in human occipital cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(4), 2077–2082.

Sayres, R., & Grill-Spector, K. (2008). Relating retinotopic and object-selective responses in human lateral occipital cortex. *Journal of Neurophysiology*, *100*(1), 249–267.

Schwartz, E. L., Desimone, R., Albright, T. D., & Gross, C. G. (1983). Shape recognition and inferior temporal neurons. *Proceedings of the National Academy of Sciences of the United States of America*, *80*(18), 5776–5778.

Schwarzlose, R. F., Swisher, J. D., Dang, S., & Kanwisher, N. (2008). The distribution of category and location information across object-selective regions of visual cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(11), 4447–4452.

Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., et al. (1995). Borders of multiple visual areas in human revealed by functional magnetic resonance imaging. *Science*, *268*, 889–893.

Sigala, N., & Logothetis, N. K. (2002). Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature*, *415*(6869), 318–320.

Swisher, J. D., Gatenby, J. C., Gore, J. C., Wolfe, B. A., Moon, C. H., Kim, S. G., et al. (2010). Multiscale pattern analysis of orientation-selective activity in the primary visual cortex. *Journal of Neuroscience*, *30*(1), 325–330.

Tootell, R. B. H., Devaney, K. J., Young, J. C., Postelnicu, G., Rajimehr, R., & Ungerleider, L. G. (2008). fMRI mapping of a morphed continuum of 3D shapes within inferior temporal cortex. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(9), 3605–3609.

Tootell, R. B. H., Mendola, J. D., Hadjikhani, N., Ledden, P. J., Liu, A. K., Reppas, J. B., et al. (1997). Functional analysis of V3A and related areas in human visual cortex. *Journal of Neuroscience*, *17*(18), 7060–7078.

Torralba, A., & Oliva, A. (2003). Statistics of natural image categories. *Network (Bristol, England)*, *14*(3), 391–412.

Troiani, V., Stigliani, A., Smith, M. E., & Epstein, R. A. (2014). Multiple object properties drive scene-selective regions. *Cerebral Cortex*, *24*(4), 883–897.

Tsao, D. Y., Freiwald, W. A., Knutsen, T. A., Mandeville, J. B., & Tootell, R. B. H. (2003). The representation of faces and objects in macaque cerebral cortex. *Nature Neuroscience*, *6*, 989–995.

Tsao, D. Y., Moeller, S., & Freiwald, W. A. (2008). Comparing face patch systems in macaques and humans. *Proceedings of the National Academy of Sciences of the United States of America*, *105*(49), 19514–19519.

Vogels, R., Biederman, I., Bar, M., & Lorincz, A. (2001). Inferior temporal neurons show greater sensitivity to nonaccidental than to metric shape differences. *Journal of Cognitive Neuroscience*, *13*(4), 444–453.

Vogels, R., & Orban, G. A. (1985). The effect of practice on the oblique effect in line orientation judgments. *Vision Research*, *25*(11), 1679–1687.

Weiner, K. S., & Grill-Spector, K. (2011). Not one extrastriate body area: Using anatomical landmarks, hMT+, and visual field maps to parcellate limb-selective activations in human lateral occipitotemporal cortex. *NeuroImage*, *56*(4), 2183–2199.

Zeidman, P., Mullally, S. L., Schwarzkopf, D. S., & Maguire, E. A. (2012). Exploring the parahippocampal cortex response to high and low spatial frequency spaces. *Neuroreport*, *23*(8), 503–507.