Representations of 3D visual space in human cortex: Population receptive field models of position-in-depth

Julie Golomb (golomb.9@osu.edu)
Department of Psychology, The Ohio State University, 1827 Neil Ave
Columbus, OH 43212 USA

Abstract:

We live in a three-dimensional world, but most studies of human visual cortex focus on 2D visual representations. The third dimension – depth – is critical for perception and behavior, yet we know relatively little about if/how position-in-depth is represented topographically in the brain, and importantly how it interacts with the well-established 2D spatial maps. We recently revealed that visual cortex gradually transitions from 2D-dominant representations to balanced 3D (2D plus depth) representations along the visual hierarchy (Finlayson, Zhang, & Golomb, 2017). Here, we ask whether this depth information is spatially organized into topographic maps, akin to 2D retinotopic maps. We employed the population receptive field modeling technique (pRF: Dumoulin & Wandell, 2008) to estimate each voxel's preferred position-in-depth and depth tuning function. Subjects viewed two different types of 3D stimuli in the scanner: depth from disparity (while wearing red/green anaglyph glasses) or depth from relative motion. Depth maps were highly reliable within a subject but demonstrated considerable across-subject variability. Yet, nearly all subjects exhibited a systematic "map-like" progression of depth-from-disparity in the vicinity of the transverse occipital sulcus. Such "depth-otopic" maps represent a novel advance carrying exciting theoretical and methodological implications for our understanding of how the brain represents spatial information.

Keywords: human visual cortex; retinotopic mapping; population receptive field models; spatial representations; depth from disparity; visual perception

Introduction

To catch a ball, reach for a cup of coffee, or find a friend in the crowd, we need to first determine the object's location. Visual information is initially recorded as precise 2D location on the retina, and much is known about how this 2D spatial information is preserved and coded in primary visual cortex and beyond (Grill-Spector and Malach, 2004). In order to perceive objects in depth, 2D retinal information must be transformed into 3D percepts, using cues such as binocular disparity, perspective, shading, and motion parallax. Binocular disparity and/or depth-sensitive responses have been reported in several visual regions in macaques and humans (Tsao et al., 2003; Ban et al., 2012; Neri et al., 2004; Preston et al., 2008; Welchman et al., 2005). However, the nature of spatial position-in-depth representations remains less explored, particularly how depth interacts with 2D spatial information to form 3D spatial representations.

A recent study from our group used fMRI multivoxel pattern analysis (MVPA) to investigate how 3D spatial information is represented throughout human visual cortex. The results revealed a widespread, systematic transition throughout visual cortex, where early visual areas primarily represented 2D location information, while position-in-depth information gradually increased along the hierarchy (Finlayson, Zhang, & Golomb, 2017). The presence of depth location information alongside 2D location in higher-level visual areas raises interesting possibilities about representations of 3D space in these regions. But it leaves open a critical question: is the depth information spatially organized into "maps" like the 2D information?

In the current study we developed two novel "depth-otopic" mapping paradigms, modifying traditional 2D phase-encoded (ring/wedge: Engel et. al., 1994; Sereno et. al., 1995) and population receptive field modeling (pRF: Dumoulin & Wandell, 2008) techniques. The traditional versions of these techniques have been used to demonstrate that human visual cortex is organized into several topographic maps of 2D spatial location (Silver and Kastner, 2009; Wandell et al., 2007). Here we modify them to probe for potential maps of depth location. We focus here on the results from the more robust pRF technique, which allows us to model each voxel's response as a tuning function, giving us information about both individual voxels' preferred depth and sensitivity (tuning width).

Methods

For the main experiment (depth from disparity), subjects viewed 3D stimuli in the scanner while wearing red/green anaglyph glasses. Depth was defined in arcmin relative to the central fixation plane (zero disparity). Positive disparity values indicate the stimulus appeared to be located in front of the central plane

(closer to the subject), and negative disparity values represent "back" stimuli (behind the central plane; farther from the subject).

A small fixation patch was located at the central plane (middle depth), and ground and ceiling frames (also always at the central depth plane) flanked the stimuli to provide additional depth references. The main stimulus was a large square (12° x 12°) composed of a random dot motion (RDM) stimulus. All dots were presented at the same depth and moved only in 2D. The dots reversed contrast from black to white every 500ms. The entire RDM stimulus changed in depth plane according to a set sequence (optimized for pRF modeling). Depth changed in steps of 3 arcmin every 4 sec, changing gradually from front (+18 arcmin) to middle (0 arcmin), middle to front, back (-18 arcmin) to middle, or middle to back. Blank gaps of 12s separated each sequence. The task was to fixate on the central depth patch and perform a fixation dot-dimming detection task while passively viewing the RDM stimulus.

Subjects were scanned in 8 runs of the pRF depth mapping paradigm on a Siemens 3T Prisma scanner, with a 32 channel coil and EPI sequences with the following parameters: 2s TR, 28ms TE, 2x2x2 mm³ resolution with coverage of occipital, posterior temporal, and posterior parietal cortex. Standard preprocessing steps were performed in SPM12 with Freesurfer used for cortical inflation. No spatial smoothing was used.

Population receptive field (pRF) modeling analyses were conducted using custom Matlab code borrowing from the SamSrf (D.S. Schwarzkopf) and analyzePRF (Kay et al, 2013) toolboxes. We employed a one-dimensional Gaussian tuning curve model (e.g. Harvey et al, 2013). Predicted fMRI timecourses were generated for different combinations of model parameters μ (preferred depth) and σ (tuning width) by modeling the hypothetical pRF response timecourses to the viewed stimulus sequence and convolving this response with a standard hemodynamic response function (HRF). The predicted timecourses were compared to the observed BOLD timecourse for each voxel, and the best-fitting model for that voxel was selected (Figure 1). Model parameters μ and σ were projected onto the cortical surface for each subject, thresholded by the goodness of fit value r².

Results

One subject was run in multiple sessions 6 months apart to assess within-subject reliability. Voxel-wise correlations between the two sessions were high for all parameters: model fit r²: r=.94; preferred depth μ: r=.59; tuning width σ: r=.69.
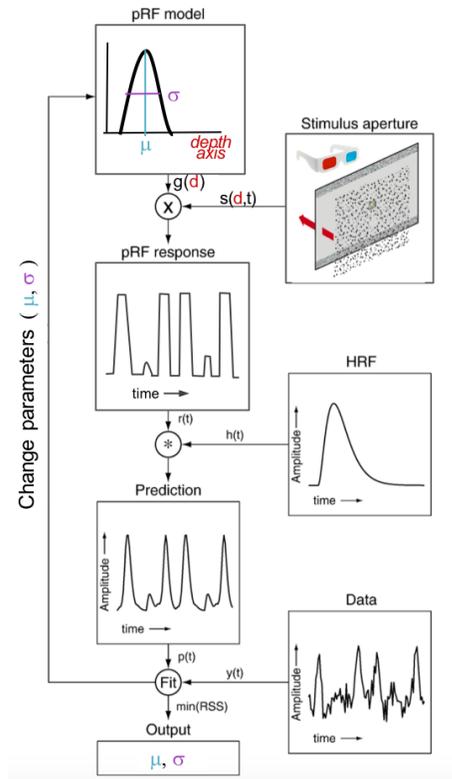

Figure 1: pRF modeling procedure. See Methods.

Voxels with strong model fits were found throughout most of visual cortex. However, most of the voxels in early visual areas exhibited very broad tuning widths, such that the tuning curves were essentially flat across all tested depths. This is consistent with the lack of decoding of depth location in early visual areas found in our MVPA study (Finlayson et al, 2017).

Among the voxels with narrower tuning curves, the precise distribution of preferred depths varied considerably across subjects. However, nearly all subjects had a cluster of depth-sensitive voxels in the vicinity of the transverse occipital sulcus (TOS, anatomically defined) that exhibited a clear map-like progression of preferred depth (see Figure 2).
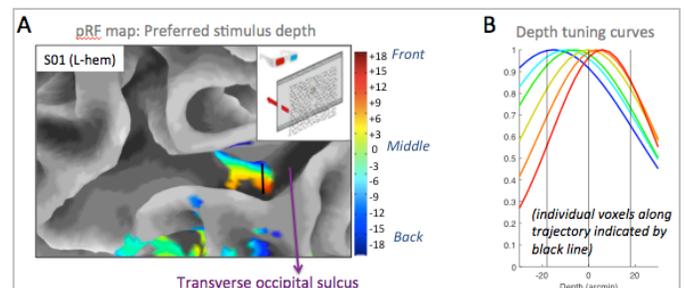

Figure 2: pRF data from sample subject illustrating (A) preferred stimulus depth and (B) depth tuning curves. Map thresholded at r > .25.

Subsets of subjects were also run in experiments comparing maps of absolute versus relative depth, and depth from binocular disparity versus depth from relative motion. The most reliable TOS maps were found for relative depth from disparity.

Conclusion

The pRF modeling technique has been previously used to investigate representations of 2D spatial location, visual field abnormalities, object size, and numerosity (e.g., Dumoulin & Wandell, 2008; Harvey et al, 2013; Baseler et al, 2011). Here we demonstrate a novel application of this technique to investigate the fundamental question of how 3D spatial information is represented in human visual cortex. The results reveal that not only can depth location be decoded from several later visual areas, but some of these areas appear to show a map-like organization of depth. In particular, we identify a putative depth map in the transverse occipital sulcus. This anatomical region overlaps with visual areas V3A, V3B, V7, and V3d: many of these regions have been previously implicated in studies of depth processing (Tsao et al, 2003). These regions also contain known retinotopic maps of 2D space (Grill-Spector & Malach, 2004). Thus, an important direction will be investigating how the discovery of these new depth maps interacts with the underlying 2D maps to support perception of 3D space.

Acknowledgments

References

Ban, H., Preston, T.J., Meeson, A., & Welchman, A.E. (2012). The integration of motion and disparity cues to depth in dorsal visual cortex. *Nature Neuroscience* 15 (4), 636–643.

Baseler, H. A., Gouws, A., Haak, K. V., Racey, C., Crossland, M. D., Tufail, A., ... & Morland, A. B.. (2011). Largescale remapping of visual cortex is absent in adult humans with macular degeneration. *Nature Neuroscience*, 14(5), 649–655.

Dumoulin, S. O., & Wandell, B. A. (2008). Population receptive field estimates in human visual cortex. *Neuroimage*, *39*(2), 647-660.

Engel, S.A., Rumelhart, D.E., Wandell, B.A., Lee, A.T., Glover, G. H., Chichilnisky, E. J., & Shadlen, M. N., (1994). fMRI of human visual cortex. *Nature*.

Finlayson, N. J., Zhang, X., & Golomb, J. D. (2017). Differential patterns of 2D location versus depth decoding along the visual hierarchy. *NeuroImage*, *147*, 507-516.

Grill-Spector, K. & Malach, R., (2004). The Human Visual Cortex. *Annual Review of Neuroscience* 27, 649–677.

Harvey, B. M., Klein, B. P., Petridou, N., & Dumoulin, S. O. (2013). Topographic representation of numerosity in the human parietal cortex. *Science*, *341*(6150), 1123-1126.

Kay, K.N., Winawer, J., Mezer, A., & Wandell, B.A. (2013). Compressive spatial summation in human visual cortex. *Journal of Neurophysiology.*

Neri, P., Bridge, H., & Heeger, D.J. (2004). Stereoscopic Processing of Absolute and Relative Disparity in Human Visual Cortex. *Journal of Neurophysiology* 92 (3), 1880–1891.

Preston, T.J., Li, S., Kourtzi, Z., & Welchman, A.E. (2008). Multivoxel Pattern Selectivity for Perceptually Relevant Binocular Disparities in the Human Brain. *The Journal of Neuroscience* 28 (44), 11315–11327.

Sereno, M.I., Dale, A.M., Reppas, J.B., Kwong, K.K., Belliveau, J. W., Brady, T. J., ... & Tootell, R. B. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science* 268 (5212), 889–893.

Silver, M.A. & Kastner, S. (2009). Topographic maps in human frontal and parietal cortex. *Trends in Cognitive Sciences* 13 (11), 488–495.

Tsao, D.Y., Vanduffel, W., Sasaki, Y., Fize, D., Knutsen, T. A., Mandeville, J. B., ... & Livingstone, M. S. (2003). Stereopsis activates V3A and caudal intraparietal areas in macaques and humans. *Neuron* 39, 555–568.

Wandell, B.A., Dumoulin, S.O., & Brewer, A.A. (2007) Visual field maps in human cortex. *Neuron* 56 (2), 366–383.

Welchman, A.E., Deubelius, A., Conrad, V., Bülthoff, H.H., & Kourtzi, Z. (2005). 3D shape perception from combined depth cues in human visual cortex. *Nature Neuroscience* 8 (6), 820–827.