

# SCIENTIFIC REPORTS



OPEN

## Image Segmentation Based on Relative Motion and Relative Disparity Cues in Topographically Organized Areas of Human Visual Cortex

Peter J. Kohler<sup>1</sup>, Benoit R. Cottureau<sup>2,3</sup> & Anthony M. Norcia<sup>1</sup>

The borders between objects and their backgrounds create discontinuities in image feature maps that can be used to recover object shape. Here we used functional magnetic resonance imaging to identify cortical areas that encode two of the most important image segmentation cues: relative motion and relative disparity. Relative motion and disparity cues were isolated by defining a central 2-degree disk using random-dot kinematograms and stereograms, respectively. For motion, the disk elicited retinotopically organized activations starting in V1 and extending through V2 and V3. In the surrounding region, we observed phase-inverted activations indicative of suppression, extending out to at least 6 degrees of retinal eccentricity. For disparity, disk activations were only found in V3, while suppression was observed in all early visual areas. Outside of early visual cortex, several areas were sensitive to both types of cues, most notably LO1, LO2 and V3B, making them additional candidate areas for motion- and disparity-cue combination. Adding an orthogonal task at fixation did not diminish these effects, and in fact led to small but measurable disk activations in V1 and V2 for disparity. The overall pattern of extra-striate activations is consistent with recent three-stream models of cortical organization.

The boundaries between objects and their background give rise to discontinuities in multiple feature maps. The role of orientation cues in the early stages of figure-ground processing has been widely studied with both single-unit recordings and functional MRI<sup>1–8</sup>, but the contribution of two strongly related parallax cues – relative motion and relative disparity – has been less well studied. Relative motion and disparity cues each contribute independently to image segmentation and the perception of shape and the two cues can be combined to disambiguate 3D object and scene structure<sup>9–11</sup>. As for orientation cues, the processing of relative motion and disparity information requires a spatial comparison of local estimates. The degree to which spatial pooling occurs at the same cortical processing stage regardless of cue, and whether the visual system applies common computational mechanisms for different cues, remains unresolved. These questions are fundamental for understanding how cue extraction in early visual areas contribute to global scene segmentation and the early stages of object recognition.

Cells in V1 of the macaque are sensitive to local motion information and sensitivity to image discontinuities created by relative motion has been observed at both early and later visual processing stages. Single unit responses to relative motion information are robust in primate V1 and V2<sup>12–17</sup>, as well as in MT<sup>18–21</sup> and IT<sup>22</sup>, although selectivity for the orientation of a motion-defined boundary may not arise until V2<sup>17,23</sup>. In humans, functional magnetic resonance imaging (fMRI) studies<sup>24,25</sup> identified an area in dorsomedial occipital cortex, originally termed the kinetic occipital area (KO), that was activated by relative motion in texture-defined bars. However, other human fMRI studies at that time found that this stimulus also produces activations in V1, V2, V3 and in the hMT+ complex that includes the homologue of macaque MT<sup>26–28</sup>. Later work made more extensive

<sup>1</sup>Department of Psychology, Stanford University, Jordan Hall, Building 420, 450 Serra Mall, Stanford, CA, 94305, USA. <sup>2</sup>Université de Toulouse, Centre de Recherche Cerveau et Cognition, Toulouse, France. <sup>3</sup>Centre National de la Recherche Scientifique, Toulouse Cedex, France. Correspondence and requests for materials should be addressed to P.J.K. (email: [pjkohler@stanford.edu](mailto:pjkohler@stanford.edu))

measurements in topographically organized visual areas and used fMRI adaptation to identify selectivity for the orientation of motion boundaries in areas V3A, V3B, LO1, LO2 and V7<sup>29</sup>, which partially overlapped with functionally-defined area KO. Thus, while evidence for relative motion processing as early as V1 has been widely reported in the macaque single-unit literature, evidence from human fMRI is less compelling.

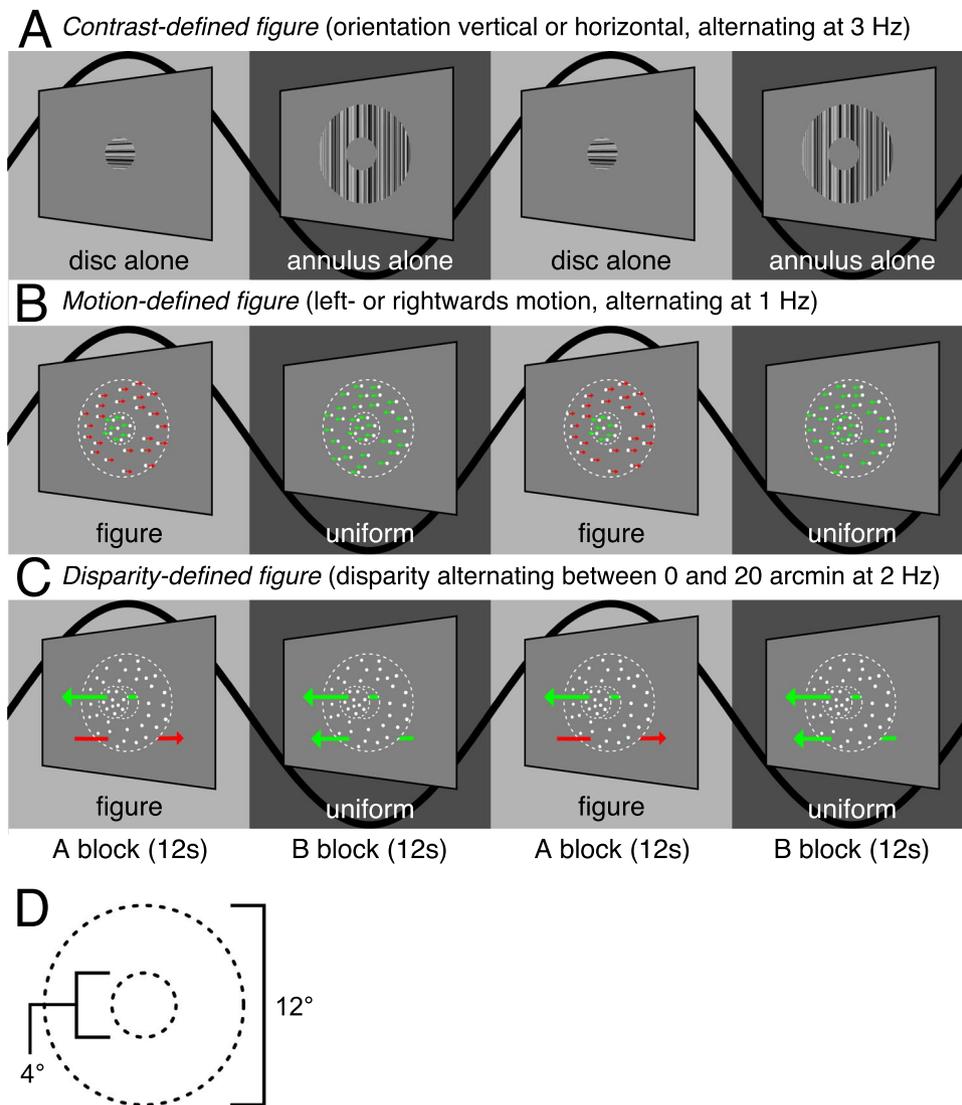
Studies in macaque suggest that image discontinuities generated by relative disparity are not encoded before V2<sup>30–32</sup>. Sensitivity to disparity discontinuities has also been found in V3<sup>33</sup>, V4<sup>34,35</sup>, IT<sup>36</sup> and, depending on the precise definition of relative disparity, in MT<sup>37</sup>. In humans, the role of early visual areas V1–V3 in processing disparity discontinuities is again less well established. Several studies have compared fMRI BOLD responses to displays with multiple disparity planes to responses produced by a single depth-plane, and found dominant activations in V3A<sup>38–40</sup>. Because these studies contrasted non-zero disparity and zero disparity displays, the effects could be driven by tuning for absolute or relative disparity, or for both (see also<sup>41</sup>). This confound was addressed by Neri and colleagues<sup>42</sup>, who used fMRI adaptation to measure responses to both absolute and relative disparity separately. They found that areas V4 and V8 (and to a lesser degree early visual areas V1, V2 and V3) showed adaptation effects for both relative and absolute disparity, but that adaptation was only present for absolute disparity in V3A, MT and V7. A more recent study of cue combination reported reliable classification of depth-sign based on motion-defined accretion and deletion cues in V1, V2, and V3 as well as in a wide range of topographically organized areas outside early visual cortex. By contrast, disparity-defined depth sign could only be robustly decoded outside early visual cortex, especially in V3B/KO<sup>9</sup>.

Global figure-ground segmentation is often modeled as relying on a combination of border discontinuity and surface-based processes<sup>43–46</sup>. Local or “first-order” encoding of both motion and disparity cues can be modelled using similar energy-like computations executed within the classical receptive field of cells in V1<sup>47,48</sup>, but local computations are insufficient to define “second-order” features such as motion- or disparity-defined edges or surfaces. The specific nature of these representations and the stage of the human visual processing hierarchy at which they are first extracted, remains poorly understood. An fMRI study of relative motion processing<sup>26</sup> suggested that local motion measurements are compared as early as V1, where BOLD activations were found to be concentrated at the retinotopic location of the border in figure-ground displays. No studies have examined whether disparity cues can drive a similar topographically specific representation. Disparity sensitivity, even at the local level, involves a prior combination of the inputs of the two eyes and thus relative disparity processing may require additional processing.

Whether early retinotopic visual areas support border- or surface-based representation of objects is differentiable in surface-based analyses of BOLD signals because border and surface-based representations will generate different retinotopically organized response profiles. We therefore compared the detailed topographic organization of relative motion and disparity responses in V1, V2 and V3. We also measured motion- and disparity-defined form sensitivity in 15 topographically organized visual areas outside early visual cortex. The latter analysis is motivated by previous results showing that figure-ground stimulus configurations such as the disk-annulus configuration are also expected to drive global shape processes and thus activate object-sensitive cortical areas<sup>49–51</sup>. Our disk-annulus stimulus configurations were presented using an experimental design that controls for the contribution of absolute motion and disparity signals to the measured responses. To probe the possible influence of attention-related feedback<sup>1</sup> on motion- and disparity-defined form processing, we had participants either passively view the stimuli or perform an orthogonal task at fixation, in separate experiments. We find that motion and disparity-defined form processing was not influenced by the orthogonal task. Overall, areas that respond to relative motion also respond to relative disparity, but representations are more robust for motion than for disparity, especially in early visual areas. The raw data from Experiments 1 and 2 are available at <http://openneuro.org/datasets/ds001984> and <http://openneuro.org/datasets/ds001978>, respectively.

## Results

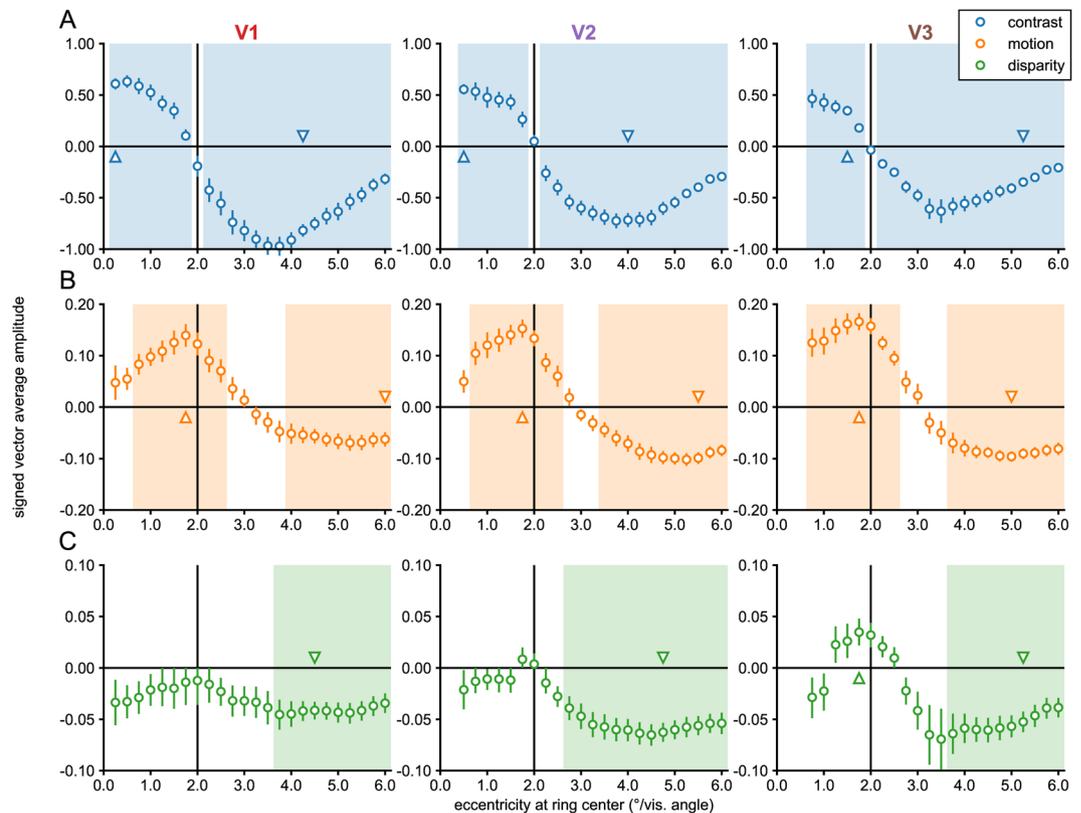
**Eccentricity analysis in early visual cortex.** We begin by describing results from early visual areas V1, V2 and V3, each of which have strong retinotopic representations of the visual field. To visualize and quantify the existence and retinotopic specificity of the representations of the motion and disparity-defined figures on the cortical surface, we used a retinotopy template<sup>52</sup> to define eccentricity-based sub-ROIs and then used a fMRI localizer stimulus to verify the accuracy of the template. The goal of the sub-ROI analysis was to track the responses to central (figure surface), boundary and outer (ground) regions of the stimulus across the surfaces of the early visual areas. The localizer stimulus consisted of 12 sec blocks of high contrast dynamic textures presented in a central disk region, alternating with the same textures presented in an annulus surrounding the disk (see Fig. 1A). We estimated the response to the stimulus as the signed average amplitude values over the period of block alternations, derived by applying a Fourier transform to the ROI-average time course for each participant, computing the vector-average amplitude and phase by averaging the real and imaginary parts of the complex values at the stimulus frequency across participants, and then assigning a positive or negative sign to each amplitude based on the phase (see Methods for details). The real and imaginary values used to compute the vector-average amplitude and phase in Experiments 1 and 2 are plotted for two V3 sub-ROIs, as well as for V3B, in Supplementary Figs 1 and 2. We used a one-sample Hotelling’s  $T^2$ -test to test if the real and imaginary parts of the complex values for each ROI were significantly different from [0,0]. The results are plotted for the contrast-based localizer condition in Fig. 2A. Positive amplitudes indicate relatively stronger responses to the A block, during which the disk was presented, compared to the B block, during which the annulus was presented. Negative amplitudes indicate the opposite. This can be observed in the single-cycle averages from the sub-ROIs that had the most significant positive and negative amplitudes in the localizer condition, plotted in Fig. 3A. For convenience, we will refer to the positive amplitudes in the localizer condition as “disk responses” and the negative amplitudes as “annulus responses”.



**Figure 1.** The experimental design used in Experiment 1. The on-off responses evoked by the conditions that are captured by our Fourier analysis, are illustrated with sine waves. **(A)** The contrast condition – the orientation of both the disk and annulus alternated between vertical and horizontal at 3 Hz. Disk and annulus were presented in temporal succession. **(B)** The motion condition – dots alternated between left- and rightwards motion at 1 Hz. When these direction changes were quadrature phase shifted between the disk and annulus regions, a figure percept was evoked, when alternations were in phase, there was no such percept. **(C)** The disparity condition – the dots alternated between the fixation plane (0 arcmin) and a position behind fixation (uncrossed disparity, 20 arcmin) at 2 Hz. Anti-phase movement of the disk and annulus regions led to a figure percept, while in phase movement generated no such percept. The white dotted line in B and C indicates the extent of the disk and annulus used in A. Note that the motion and disparity conditions were identical to those used in Experiment 2, except for the introduction of an RSVP task at fixation, which directed attention away from the stimulus. **(D)** Schematic outlining the size of the disk and annulus region, respectively, which was shared across all conditions.

In areas of cortex defined by the template as responsive to eccentricities covered by the disk, we see high amplitude disk responses. At eccentricities defined as near the disk boundary (2° radius), the phase sign reverses and we begin to see annulus responses. There are significant responses at all eccentricities covered by the stimulus, except for the sign reversal region where voxels responding to the disk and the annulus, 180° out of phase, are likely mixed together. The pattern of results indicates that the template-based procedure allows us to accurately localize the eccentricity of the contrast-defined boundary in retinotopic cortex.

While the contrast condition alternated a disk and an annulus, the motion and disparity conditions alternated figure-ground and uniform field configurations (see Fig. 1). The dots updated dynamically at 20 Hz throughout each block, and in the motion condition, they were temporally correlated with a long lifetime. In the disparity condition, they were temporally uncorrelated across updates, and thus did not generate monocular cues to form. In the motion condition, dots alternated between left- and rightwards movement at 1 Hz, while in the disparity

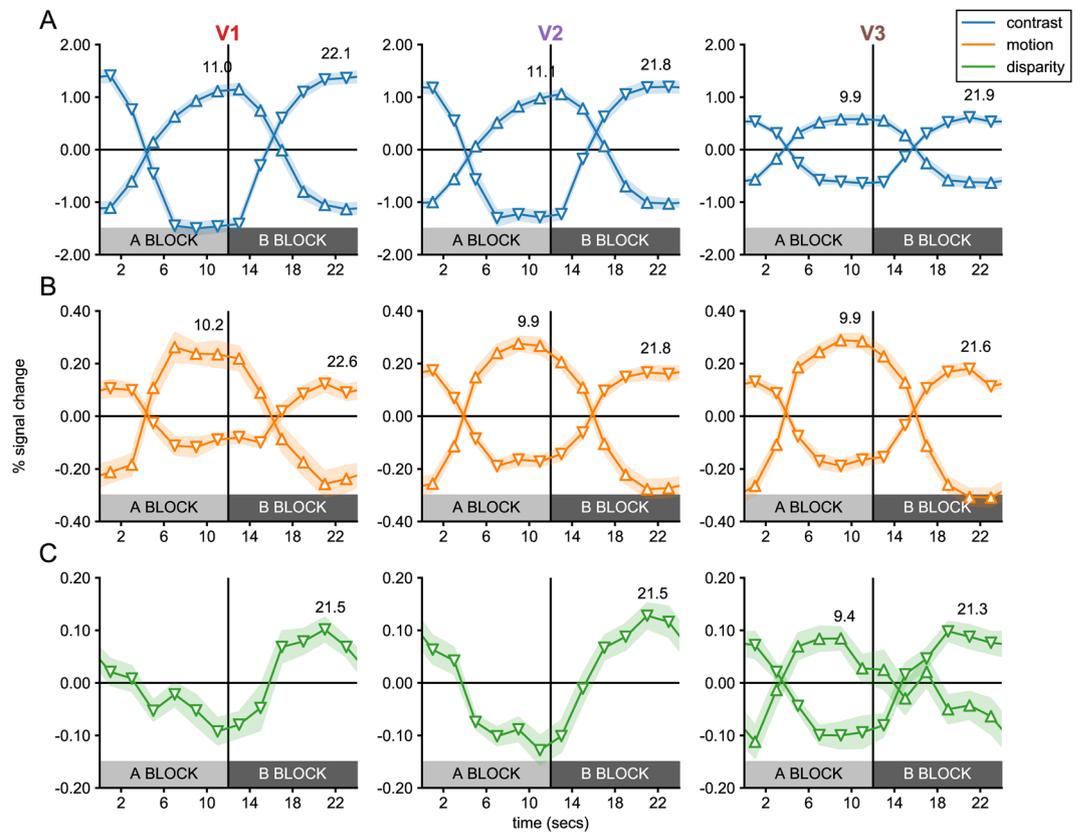


**Figure 2.** Eccentricity analysis of V1–V3 in Experiment 1. Signed vector-average amplitudes for the contrast (A), motion (B) and disparity (C) conditions within 24 sub-ROIs, defined based on a retinotopy template as centered on eccentricities spaced  $0.25^\circ$  apart, each having a width of  $0.5^\circ$ , as plotted on the x-axes. Note that V2 and V3 sub-ROIs centered at low eccentricities ( $0.25^\circ$  for V2;  $0.25^\circ$  and  $0.5^\circ$  for V3) were not analyzed and plotted (see Methods). The shaded areas on the plots indicate condition  $\times$  sub-ROI combinations that were significant at  $\alpha = 0.05$ . The positively and negatively signed sub-ROIs that had the lowest  $p$ -values for each combination have been indicated with up and down-arrows, and single cycle averages have been plotted for these sub-ROIs in Fig. 3. For disparity, no V1 sub-ROIs were given a positive sign, and the two positively signed sub-ROIs in V2 were far from significance, so the positively signed cycle average for disparity was only plotted for V3, where the lowest positively signed  $p$ -value was close to significance ( $p = 0.06$ ). The real and imaginary values used to compute the signed vector-average amplitudes have been plotted for the highlighted negatively and positively signed V3 sub-ROIs in Supplementary Fig. 1. The color of the text for the ROI names matches the ROI colors plotted on the inflated cortical surface reconstruction in Fig. 6.

condition they alternated between the fixation plane ( $0$  arcmin), and a position behind fixation (uncrossed disparity,  $20$  arcmin) at  $2$  Hz. For both conditions, the disk and annulus region alternated in anti-phase during the A block, and in phase during the B block. Importantly, absolute motion and absolute disparity were alternating during both A and B blocks, but only the A block generated relative motion and disparity cues that give rise to a figure boundary. The design and corresponding Fourier analysis yields the differential response between the blocks and thus allows us to isolate responses driven by the presence of relative motion/disparity discontinuity, the segmented figure surface or some combination of the two, from responses driven by absolute motion and disparity.

The overall response amplitudes in the motion and disparity conditions were 5–10 times weaker than in the contrast condition, but there were nonetheless significant activations at multiple eccentricities. For both motion and disparity, positively signed amplitudes indicate stronger responses to the A block in which the figure and surround were segmented, while negatively signed amplitudes indicate stronger responses to the B block in which a uniform field was perceived. As we discuss the results, we will use the terms “figure responses” and “uniform responses” to refer to positively and negatively signed amplitudes, respectively. Note that these terms are meant to strictly refer to the phase sign of the response relative to the A and B blocks, and not as an interpretation of the data. As noted in the Introduction, figure responses can be driven by enhancement of the figure boundary or surface, or by a mixture of these two forms of enhancement. Uniform response indicate that when the figure is present, responses are weaker than when the figure is not present. Uniform responses that are measured in the region outside the figure are likely due to suppression of the surround when the figure is present, rather than surround-exclusive enhancement of responses when the field is uniform.

The motion condition (see Fig. 2B), produced figure responses at the disk boundary in all early visual areas, which persisted inside the disk region and could also be observed  $\sim 1^\circ$  outside the disk. Surprisingly, we also saw uniform responses at eccentricities beginning  $\sim 3^\circ$ , which were significant in all visual areas. In the disparity

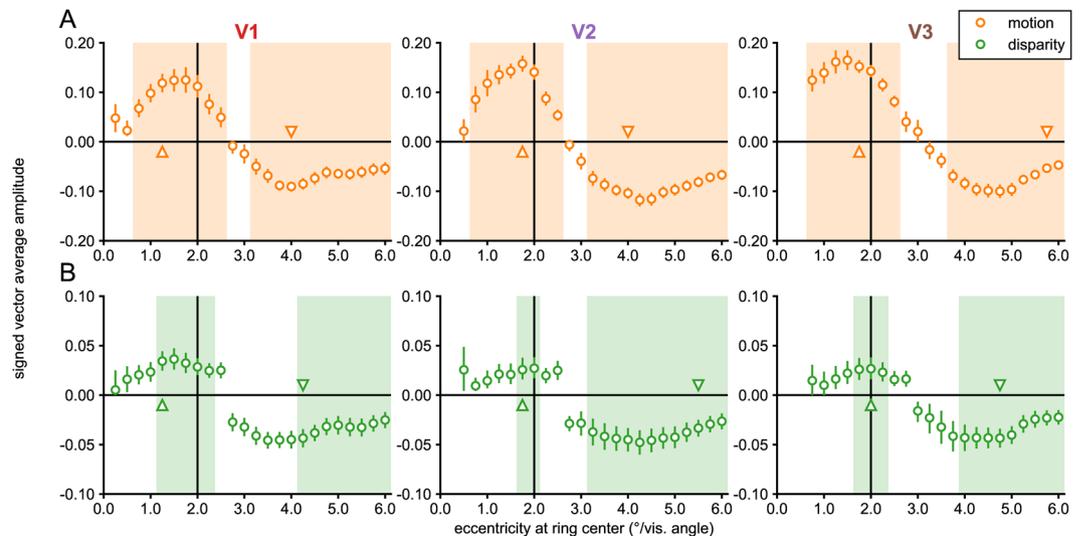


**Figure 3.** Single cycle averages for V1, V2 and V3 sub-ROIs in Experiment 1. Single cycle averages were computed from the most significant positively and negatively signed sub-ROIs in each visual cortex ROI following the color and marker convention of Fig. 2. The shaded region around each plot indicates the standard error of the mean, across participants. The delay of the peak response in seconds, estimated based on the phase, is plotted above each cycle average. Note that cycle averages have been plotted at the middle of each 2-second TR. The plot clearly demonstrates that positively-signed sub-ROIs (up-arrows) respond preferentially to the A block, while negatively signed ROIs respond preferentially to the B block (down-arrows).

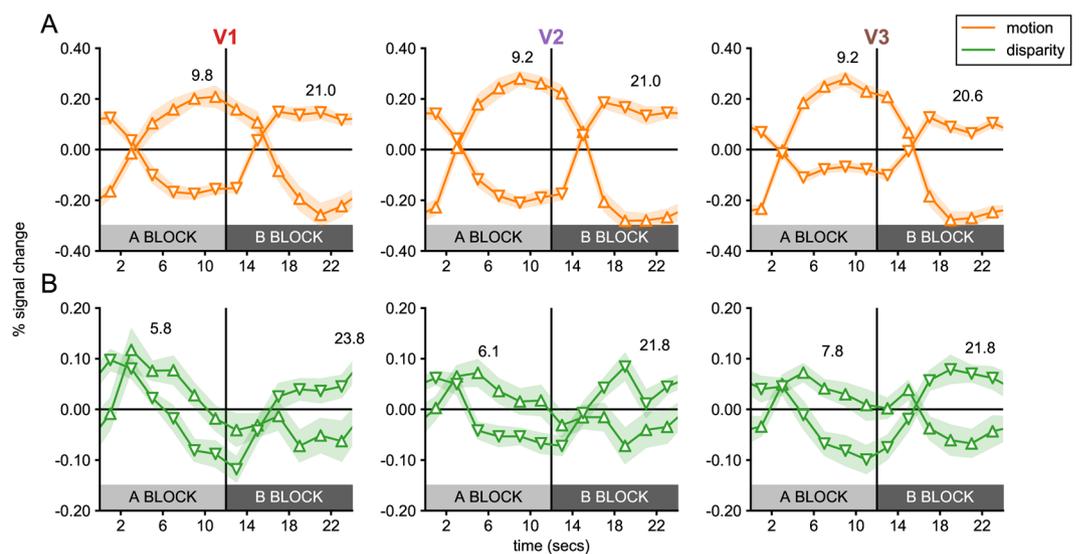
condition (see Fig. 2C), we did not measure significant figure responses in V1 or V2, and in fact hardly ever observed positively signed amplitudes, according to our phase-derived definition (see Methods). No V1 sub-ROIs had positive amplitudes, and the only two V2 sub-ROIs that did, were far from significant ( $p$ 's  $> 0.7$ ). By contrast, a region of six V3 sub-ROIs centered on the boundary and extending  $\sim 0.5^\circ$  inside and outside the boundary were given positive signs. The two sub-ROIs at and immediately inside the boundary approached significance ( $p$ 's  $< 0.08$ ; smallest  $p = 0.06$ ). We also observed significant uniform responses in all early visual areas. Single cycle averages for the figure and uniform responses are plotted for motion and disparity in Fig. 3B,C. Figure responses had similar shapes and estimated delays to the disk responses in the localizer conditions, while uniform responses in both conditions were similar to the annulus responses.

These results suggest an interesting dichotomy: Among the early visual areas, we only see figure responses for relative disparity in V3, while all three early visual areas have uniform responses. As mentioned above, the uniform responses measured for motion and disparity are likely the result of suppression when figure is present. This relative suppression of surround-region responses may be due to feedback from higher cortical areas, driven by top-down attentional selection. Spatial attention can modulate BOLD responses in all early visual areas, including V1<sup>53–55</sup>. We directly tested the hypothesis that the suppression is due to attention-driven feedback by running a second experiment.

Experiment 2 used the same parameters as the motion and disparity conditions in Experiment 1, except that attention was now directed away from the stimulus via an orthogonal task presented at fixation (see the 'Visual stimuli' section in the Methods). Response profiles were generally very similar across the two experiments. We tested this by performing independent samples Hotelling's  $T^2$ -tests comparing the data between the two experiments for each of the 69 sub-ROIs across the three visual areas (see Methods). None were significant for the disparity condition and only 3 for the motion condition, both in the V1 periphery where amplitudes were slightly larger for the passive viewing condition (in sub-ROIs centered on  $5.5^\circ$ ,  $5.75^\circ$  and  $6.0^\circ$ ; biggest  $p = 0.005$ ). As expected from the failure to detect consistent significant differences between the experiments, both motion and disparity produced uniform responses that were largely similar between the experiments. Yet, despite the results of the between-group test, the disparity condition in Experiment 2 produced significant figure responses, as measured with a one-sample test against 0, near the figure boundary in all three early visual areas (see Fig. 4), in contrast with Experiment 1 where evidence of disparity-driven figure responses was only seen in V3 (see Fig. 2).

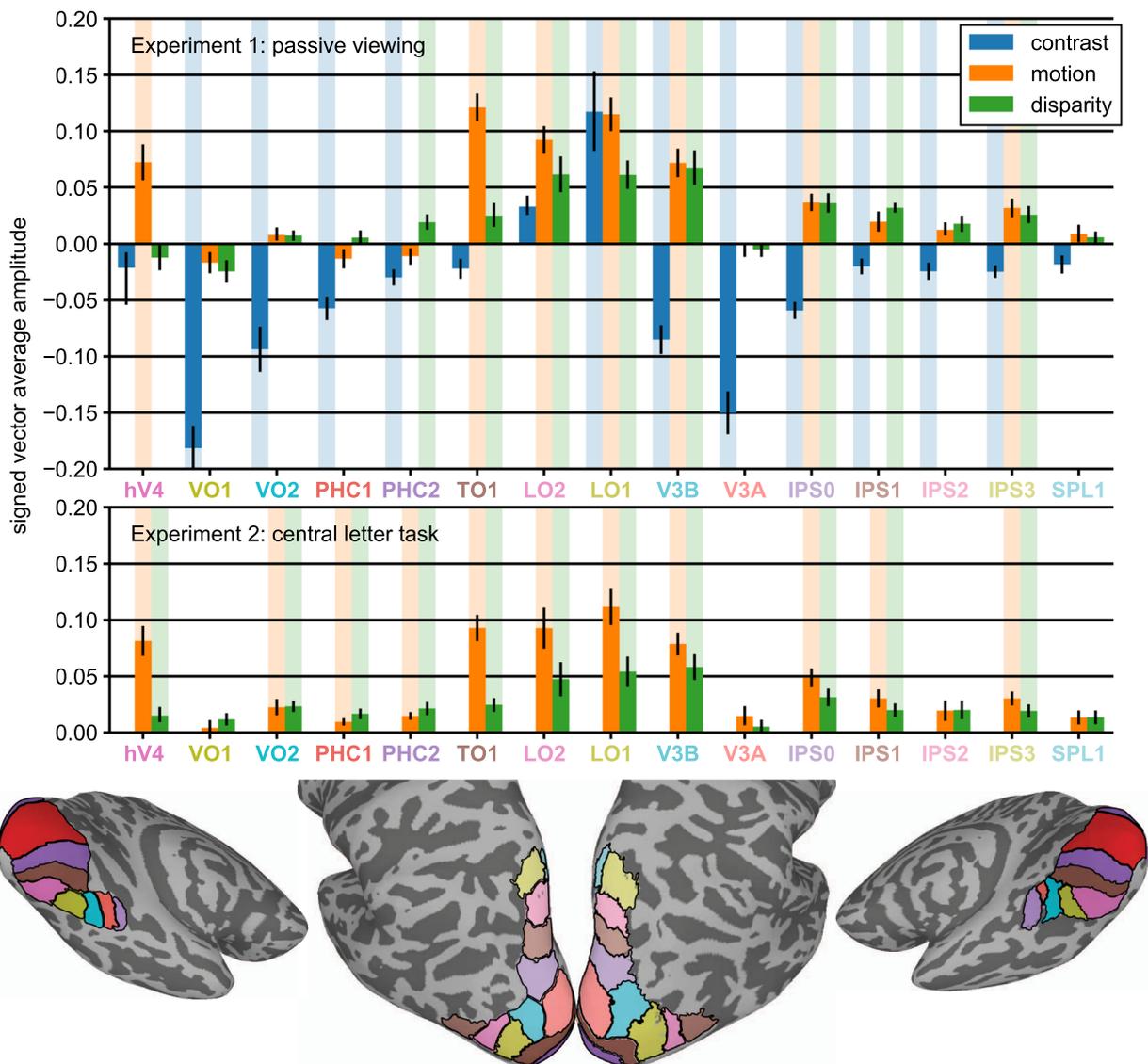


**Figure 4.** Eccentricity analysis of V1-V3 in Experiment 2. Signed vector-average amplitudes within 24 eccentricity-defined sub-ROIs, centered on eccentricities spaced  $0.25^\circ$  apart, each having a width of  $0.5^\circ$ . The shaded areas on the plots indicate condition  $\times$  sub-ROI combinations that were significant at  $\alpha = 0.05$ . Note that V2 and V3 sub-ROIs centered at low eccentricities ( $0.25^\circ$  for V2;  $0.25^\circ$  and  $0.5^\circ$  for V3) were not analyzed and plotted (see Methods). The up and down-arrows respectively indicate the positively and negatively signed sub-ROIs that had the lowest  $p$ -values for each combination. The real and imaginary values used to compute the signed vector-average amplitudes have been plotted for the highlighted negatively and positively signed V3 sub-ROIs in Supplementary Fig. 2. The color of the text for the ROI names matches the ROI colors plotted on the inflated cortical surface reconstruction in Fig. 6. The motion and disparity conditions plotted here were identical to the ones used in Experiment 1 (see Fig. 2), except attention was directed away from the stimulus.



**Figure 5.** Single cycle averages for V1, V2 and V3 sub-ROIs in Experiment 2. Single cycle averages computed for the most significant positively and negatively signed sub-ROIs in each visual cortex ROI, following the color and marker convention of Fig. 2. The shaded region around each plot indicates the standard error of the mean, across participants. The delay of the peak response, estimated based on the phase, is plotted above each cycle average. Note that cycle averages have been plotted at the middle of each 2-second TR. The plot clearly demonstrates that positively-signed sub-ROIs (up-arrows) respond preferentially to the A block, while negatively signed ROIs respond preferentially to the B block (down-arrows).

Single-cycle averages were also quite similar across the two experiments, although the delays of the peak figure response were perhaps slightly shorter for disparity in Experiment 2, compared to the motion condition and the other conditions in Experiment 1 (compare Figs 3 and 5).

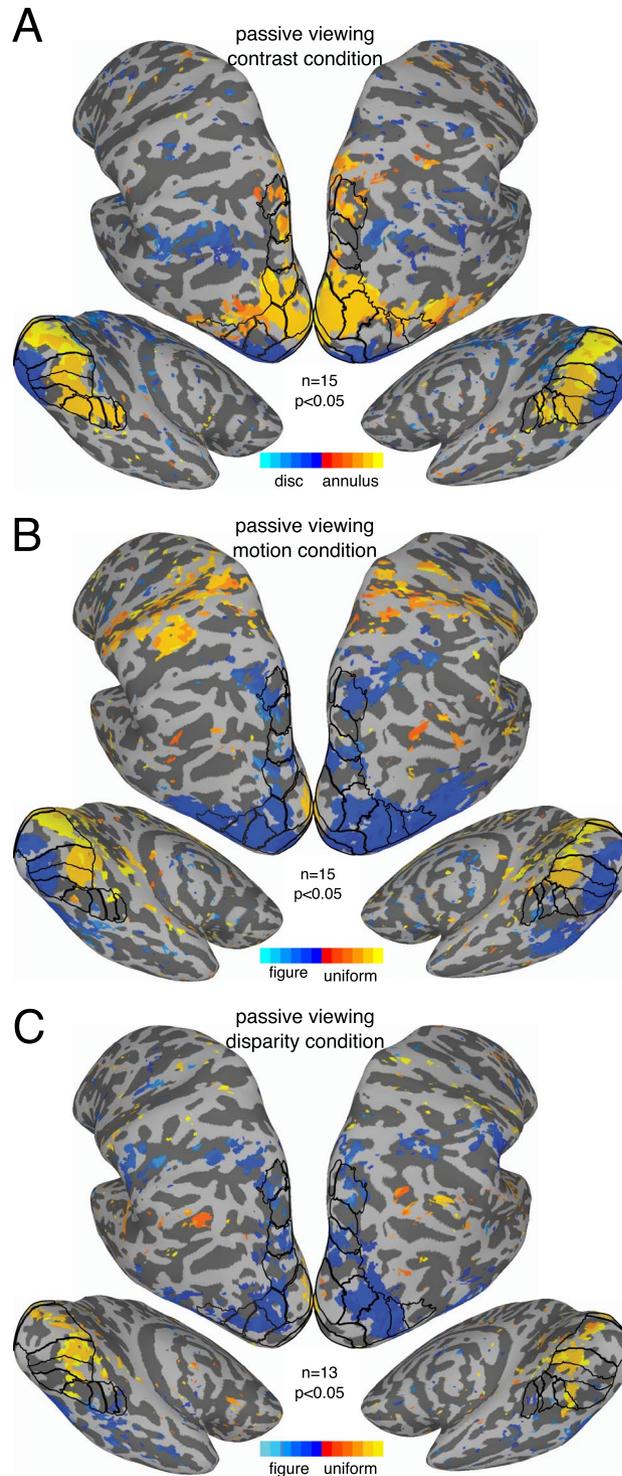


**Figure 6.** ROI results from Experiments 1 and 2. Signed vector-average amplitude within 15 topographically organized regions of interest, excluding early visual areas V1-V3, which are plotted in Figs 2–5. The contrast condition is shown in blue, the motion condition in orange, and the disparity condition in green. Shaded areas behind the bars indicate condition  $\times$  ROI combinations that were significant at  $\alpha = 0.05$ . The ROIs are shown for the both hemispheres of an example participant's inflated cortical surface reconstruction below the graph.

These results lead us to reject our hypothesis that the uniform responses we saw in the surround region in Experiment 1 are due to top-down attention-mediated suppression of the surround. Instead, the results if anything suggest that spatial attention elicited by the orthogonal task at fixation may in fact enhance disparity-driven figure responses in V1 and V2, making weak signals measurable. In V3, figure responses are present in both experiments, but the figure responses are perhaps more robust with the orthogonal task (where 3 sub-ROIs reached significance).

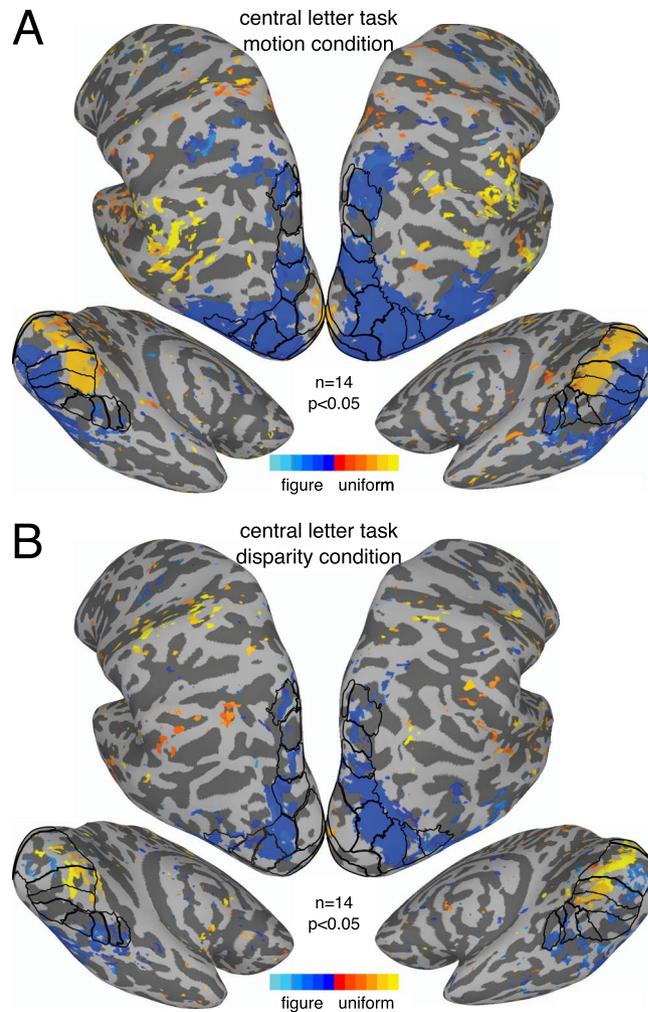
**Extended ROI-based analysis of responsivity.** The next set of analyses quantifies motion and disparity sensitivity in topographically organized visual areas beyond early visual cortex. Here the signed vector-average amplitude was computed based on the average time course across all voxels within each ROI. As for the sub-ROIs in early visual cortex, positive amplitudes indicate responses to the figure (A block), while negative amplitudes indicate stronger responses to the uniform field (B block).

In both the motion- and disparity-defined form conditions, all significant activations ( $p < 0.05$ ; indicated with shaded areas in Fig. 6) were positively signed. This is consistent either with positive-sign activations generated by the figure overcoming any negatively signed activations that may have occurred in a subset of voxels within a given ROI, or a lack of negatively signed activation. We distinguished three response patterns: areas that only had significant responses to the motion-defined figure (hV4), areas that only had significant responses to the disparity-defined figure (PHC2, IPS1) and areas that had significant responses to both (TO1, V3B, LO1, LO2,



**Figure 7.** Whole-brain results from Experiment 1. Vector-averaged phase maps, converted to seconds of delay, thresholded at  $\alpha = 0.05$ . Maps were produced through surface-based alignment procedure in which each subjects' cortical mesh was converted to a standardized mesh, which allowed for cross-subject comparisons of values at each mesh node. The vector-average phase across subjects, as well as a corresponding  $p$ -value based on both amplitude and phase, could then be computed for each mesh node. As indicated by the color bar, orange shades indicate responses to the A block, while blue shades indicate responses to the B block. The ROIs are outlined on the surface, and labeled versions can be inspected in Fig. 6.

IPS0, IPS3). Our results suggest a clear functional distinction between V3B and nearby area V3A, which has no significant responses. Motion generated stronger responses than disparity in V3 (see Figs 2 and 4), but this was less pronounced among the higher-level areas: Both V3B, IPS0 and IPS3 had nearly identical responses to the two conditions.



**Figure 8.** Whole-brain results from Experiment 2. Vector-averaged phase maps, converted to seconds of delay, thresholded at  $\alpha = 0.05$ . Maps were produced as described for Fig. 7 and in the text. As indicated by the color bar, orange shades indicate responses to the A block, while blue shades indicate responses to the B block. The ROIs are outlined on the surface, and labeled versions can be inspected in Fig. 6.

For comparison, we also plot the contrast-based localizer condition in Fig. 6. Most ROIs produce negatively signed responses, indicating that more voxels are responding to the annulus compared to the disk. LO1 and LO2 are the only exceptions to this rule, and the positively signed responses (which reach significance in LO1) are likely a result of the known preference of these areas for global shapes<sup>49</sup>. It is also worth noting that several ROIs that produce weak, non-significant responses in the motion and disparity conditions, nonetheless respond strongly in the contrast condition (VO1, VO2, PHC1, V3A), indicating that the observed variability in motion and disparity sensitivity cannot be explained by reference to general difficulties in measuring responses from specific ROIs.

**Effect of attention on disparity-defined figure activations outside of early visual cortex.** The results of Experiment 2 are shown in the bottom half of Fig. 6. Did the orthogonal task influence the responses to motion and disparity? Experiment 2 produced significant positively signed responses to motion, as measured with a one-sample test against 0, in every area that had reached significance in Experiment 1, as well as in VO2, PHC1, PHC2 and IPS1, although the independent samples test comparing the motion conditions across the two experiments only reached significance in PHC2 ( $p = 0.044$ ; all other  $p$ 's  $> 0.15$ ). For disparity, we again replicated every significant effect from Experiment 1, with additional significant effects observed in hV4, VO2, and PHC1. The independent samples test reached significance in VO1 ( $p = 0.029$ ), an area for which none of the single sample tests reached significance, and approached significance in VO2 ( $p = 0.065$ ; all other  $p$ 's  $> 0.25$ ). These results suggest that the figure responses we measured for motion- and disparity are mostly independent of attention, at least for the tasks we used here. The small effects that we do observe, mostly in ventral areas, are in the form of enhanced figure responses, similar to what we observed for disparity in V1 and V2 sub-ROIs.

**Whole brain analysis.** A surface-based alignment approach was used to visualize vector-averaged responses to the conditions in the two experiments, across all of cortex, including regions outside our set of ROIs. The results of these analyses did not deviate from the sub-ROI and whole-ROI analyses in any material way, and we will only describe them briefly. We plot the phase of the vector-averaged response, converted to seconds of delay, and thresholded by significance. Blue colors indicate responses to the A block, while orange colors indicate responses to the B block. For the contrast condition (see Fig. 7A), the reversal of the phase sign from disk responses to annulus responses between low and high eccentricities, described in the sub-ROI analysis (see Fig. 2A) can be clearly observed in early visual areas.

For motion, we see clear evidence of the figure region response in early visual cortex near the occipital pole (Fig. 7B, blue colors) and uniform responses consistent with suppression of the surround more anteriorly (orange colors). We also see figure responses that cover most dorsolateral ROIs and extend anteriorly and ventrally beyond the ROIs. The figure response is weak or absent in V3A or IPS2, but strong in nearby areas such as V3B, LO1, IPS0 and IPS3. The general pattern is similar in Experiment 2 (see Fig. 8A), although figure responses in ventral areas are perhaps slightly more pronounced.

For disparity, we again see uniform responses consistent with suppression in early visual cortex (see Fig. 7C, orange colors). It is worth noting that for both motion and disparity, in both experiments, there is little evidence of suppression outside of early visual cortex, at least within our ROIs. Figure responses can be observed in early visual cortex in Experiment 2 (see Fig. 8), although they are fairly spotty and likely benefitted from the increased sensitivity of the sub-ROI analysis (shown in Figs 2 and 4). Both experiments produce figure responses in V3B, LO1, and several IPS areas, but not in V3A. As for motion, figure responses in ventral areas are perhaps slightly stronger in Experiment 2, than in Experiment 1.

## Discussion

Figure-ground segmentation could be supported by several mechanisms that may act separately or in combination, including, as noted by Likova and Tyler<sup>56</sup>, retinotopic enhancement of the figure boundary, retinotopic enhancement of the figure surface and retinotopic suppression of the ground region. Our detailed measurements of the representation of motion- and disparity-defined forms in strongly retinotopic visual areas V1, V2 and V3 provide evidence that all three of these mechanisms are implemented in early visual cortex, at least for motion. Our measurements in retinotopic areas outside early visual cortex demonstrate that with a few important exceptions, areas that respond to one cue also respond to the other. Responses to both cues are stronger in lateral and dorsal visual areas compared to areas on the ventral surface of cortex, extending a similar pattern found in previous studies to a wider range of areas.

**Representations of the figure boundary and surface in early visual areas.** Our results provide clear evidence for enhancement of the figure boundary. For motion, we find measurable figure responses around the boundary in all early visual areas and in both experiments. For disparity, figure responses near the boundary are measurable in V3 in Experiment 1, and in all early visual areas in Experiment 2.

The evidence for surface enhancement is more ambiguous. For the motion and disparity conditions, figure responses are weaker towards the center of the figure and become non-significant at low eccentricities for V1 and V2. However, the drop off is more pronounced on the outer side of the figure boundary than on the inner side. This is especially clear for the motion condition where responses are strong across all early visual ROIs in both experiments. This asymmetry indicates that voxels with receptive fields centered inside the figure boundary are more likely to produce a figure response than those just outside the boundary. Are voxels inside the figure simply responding to the figure boundary, rather than the surface? Population receptive fields measured in voxels increase in size towards the periphery<sup>57</sup>, so voxels outside the boundary are more likely to have receptive fields that overlap with the boundary, compared to those inside. This means that boundary-exclusive responses should be more pronounced outside the figure, compared to inside, the opposite of the asymmetry we observe. So, while responses near the center of the figure become weaker compared to the contrast condition, there is clear evidence for some form of surface enhancement for motion-defined form. Disparity may also give rise to surface enhancement, but this is harder to measure because early visual cortex responses are generally less robust for disparity than for motion.

It is perhaps surprising that we see any figure responses outside the figure boundary. This likely reflects two factors, one is the size of the receptive fields that best respond to motion and disparity and the other the relative balance of more global processes that leads to retinotopically organized enhancement and suppression. The apparent size of the center and surround regions may depend on both these factors. If the underlying receptive fields are relatively large and the strength of enhancement is greater than that of suppression, then the disk representation may appear to be expanded.

Phase-inverted responses consistent with suppression of the ground region (see below) were measured across all early visual areas for both motion and disparity and this suppression is independent of the task we used. Thus, for motion-defined form we have strong evidence for boundary and figure enhancement as well as ground suppression in all three early visual areas. For disparity, we have strong evidence for boundary enhancement and ground suppression for disparity, but no conclusive evidence of figure enhancement.

**Basis of opposing sign responses.** The patterns of activation we observed in V1, V2 and V3 for motion- and disparity-defined figures each consisted of a region of positively signed activation surrounded by a negatively signed activation. The negatively signed activation in the surround could arise from suppression of responses during the A block or enhanced responses during the B block. We consider the former more plausible for several reasons. First, the center and surround region are indistinguishable in the uniform stimuli used in the B blocks, and it is thus unlikely that there would be selective enhancement of the surround. Second, an

opposing sign organization like the one we have observed has been found previously in macaque for “first-order” contrast-defined stimuli when comparing blank screen baseline activation levels to those within and adjacent to the retinotopic location of high contrast patterns<sup>58–60</sup>. Single unit and local field potential recordings in the retinotopic and adjacent representations of spatially localized, high contrast stimuli indicate suppression of baseline neural responses in regions that had negatively signed BOLD<sup>58</sup>. Our results suggest that similar suppression is occurring with our motion- and disparity-defined “second-order” contrast stimuli. Suppression of spiking activity from the ground region by the presence of a figure has been reported for orientation-defined textures in macaque V1<sup>2</sup>.

The combination of figure/surface enhancement and surround suppression may thus be a common computational strategy across several cue types, including contrast<sup>58,59</sup>, texture<sup>2</sup>, as well as the motion and disparity cues measured here. The strength of the spatial antagonism of responses in early visual areas may differ across cues, perhaps because disparity requires binocular processing, while other cues can be processed monocularly. Nonetheless, our results, in combination with existing results in the literature, suggest that there is a substantial degree of invariance in the computational strategy applied for motion, disparity, contrast and texture as early as V1. Interestingly, in our hands both enhancement and suppression are largely independent of task.

**Comparison with previous fMRI studies of figure/ground segmentation.** The center-surround configuration of our disk stimuli lends itself to the detection of alternate-sign activations because eccentricity is mapped systematically on the cortical surface in the foveal confluence-region of early visual cortex<sup>61</sup>. In the motion-domain, Reppas and colleagues<sup>26</sup> observed positive activation at the border of a disk-shaped motion-defined form, but found neither negative activation in the ground region nor positive activation within the figure region, as we see in our data. This difference may be due to an ice-berg effect in the older dataset, where only the strongest activations rise above the noise floor. A suggestion of negative BOLD activation was present in the disparity data of Parker and Bridge<sup>62</sup>, but their use of rotating wedge-shaped stimuli complicated its visualization and measurement. A study in which a second-order figure region (a bar) was defined by temporal transients<sup>56</sup> did find negative activation adjacent to the retinotopic locus of the bar, but that study did not find enhanced activation within the retinotopic representation of the figure as we did. Texture-defined figures have also been found to increase overall activation in early areas<sup>5</sup>, but the analysis did not consider figure, boundary and ground regions separately as we have done here. Finally, isolated background suppression has also been found for luminance-defined shapes, but this study did not evaluate figure enhancement effects<sup>63</sup>.

**Role of top-down influences.** Several studies have suggested that figure/ground representations in V1 depend on influences from higher-order visual areas. For example, figure enhancement for texture-defined form that is measurable in V1 in awake animals is reduced by anesthesia<sup>64</sup>. Simultaneous recordings in V1 and V4 suggest that surround-region suppression in V1 for texture-defined forms is controlled by an input from V4<sup>1,2</sup>. These previous studies motivated our use of two different behavioral tasks in our study of motion- and disparity-defined form. In one task, the participants were instructed to fixate in the center of the figure region and to pay attention to the stimuli, in the other they were given a difficult letter discrimination task to perform in which the letters overlapped the figure region. The most prominent effect of the task was that responses to the disparity-defined figure boundary became measurable in V1 and V2 when participants were performing the letter task. Although the direct comparison of Experiments 1 and 2 did not reach significance in any sub-ROIs near the figure boundary, this difference may reflect a true effect of attention. Such an effect would indicate that allocating spatial attention to a region of the visual field can upregulate weak figure processing in early visual cortex, in addition to the stimulus-independent modulation of responses previously observed<sup>53–55</sup>. This is an intriguing possibility that should be explored further.

The task had limited influence on negatively signed activations in early visual areas. It is thus possible that the suppression we observe may depend on feedback mechanisms that are relatively independent of task-related deployment of attention. Early visual areas receive direct feedback connections from a wide range of cortical areas, including superior and inferior temporal cortex<sup>65</sup>, V2 and MT (see<sup>66</sup> for review).

**Comparison with single-unit studies of boundary and figure responses in EVC.** Single-cell recordings have identified selectivity for motion-defined edges as early as V1<sup>13,15</sup>, but similar selectivity for disparity has been observed only as early as V2<sup>67</sup>. The extraction of spatial cue-relationships necessary for detecting an edge could be made over relatively small distances, while surround modulation requires longer-range spatial interactions. We find suppressed responses over the full extent of our stimulus presentation, so as far as 6° of radius and 4° from the border of the figure region. Several single-unit studies of texture segregation have explored the spatial organization of responsiveness to second-order figure-ground displays by placing the receptive field (RF) of V1 neurons at different locations within the figure, at its border or on the background region. These prior results from texture- and motion-defined form demonstrate enhanced responses to the figure and boundary region, relative to responses when the receptive field was on the background<sup>1,2,64,68</sup>, broadly consistent with our results.

Our fMRI block-design contrasted responses during periods where the field was segmented (A blocks) and periods where it was uniform (B blocks). To observe both figure enhancement and suppression of the background in the presence of a figure, it is critical to compare responses to segmented and uniform fields, and measure responses in both the figure and surround regions. Most single-unit studies have not done both. The single study of disparity-defined form found figure region enhancement in V1 measured relative to a uniform field, but no measurements were made in the surround<sup>69</sup>. The only single-unit study that has done both highlights the necessity of contrasting segmented and uniform field responses: Poort and colleagues showed explicit enhancement

of activity in the figure region and suppression of the background, using texture-defined figures<sup>2</sup>. The measured suppression was sufficiently long-range that a figure presented in one hemisphere could elicit suppression in the other<sup>2</sup>. This is consistent with the suppression we measure at locations several degrees beyond the figure/ground border.

**Figure/ground representation in higher-order visual areas.** We observed stronger responses to relative motion compared to relative disparity in early visual areas (see Fig. 2). This difference was present, but less pronounced in higher-level areas with responses being especially similar in V3B, IPS0 and IPS3, as has also been observed by Bridge and Parker<sup>62</sup>. Our results are also consistent with the ‘single-cue’ classification accuracies reported by Ban and colleagues<sup>9</sup>, which are greater for motion than disparity in early visual areas, but more comparable in higher-level areas<sup>9</sup>. In V3A, their ‘single cue’ classification accuracies are above-chance for stimuli that do not produce significant responses in our data. The classifier may be picking up tuning for absolute motion and disparity<sup>9</sup>, which was controlled for in our experiment design. In this interpretation, V3A is sensitive to absolute, but not relative, motion and disparity, while neighboring V3B is sensitive to relative motion and disparity.

Outside of early visual cortex, we found that all areas that respond to one cue respond to the other cue in at least one of the two experiments. Areas TO1, LO1, LO2, V3B, IPS0 and IPS3 were jointly responsive in both experiments, while hV4, VO2, PHC1, PHC2 and IPS1 were jointly responsive only in Experiment 2. For the latter group of ROIs, joint responsiveness thus depends on attentional enhancement of weak processing of figures defined by one of the two cues, likely driven by a figure-specific effect of spatial attention, as described above for early visual cortex. Areas that are jointly responsive to the two cues are candidate regions for cue-combination for motion and disparity information. The task-independent responsiveness found in TO1, LO1, LO2, V3B, IPS0 and IPS3 may provide a more robust foundation for combining the two cues. Our analysis does not specifically address whether cue-combination occurs in the areas we find to be responsive to both cues, but it does provide a set of candidate areas in which cue combination could possibly occur. Ban and colleagues<sup>9</sup> found evidence of integration of motion and disparity in V3B/KO, which partially overlaps with several of the areas we find to be sensitive to both cues: V3B, LO1, LO2 and V7/IPS0. This does however leave several jointly responsive areas that do not overlap with V3B/KO: TO1, IPS1, IPS3, hV4, VO2, PHC1 and PHC2, of which TO1 and IPS3 were jointly responsive in both experiments.

**Possible limitations.** As suggested by prior single-unit work, the fundamental contrast to be made is between segmented and uniform fields. Our fMRI block design allowed us to measure this contrast and compare responses during segmented A blocks and uniform B blocks. However, for the motion conditions, the A block alternated between uniform and segmented fields, while in the disparity A block the field was always segmented. This difference is unlikely to have changed the spatial pattern observed in early visual areas, but could affect the relative magnitude of motion and disparity responses across visual cortex. Resolving this second-order effect requires further study.

**Three streams.** Relative disparity processing has been associated with the “canonical macaque ventral stream” leading from V4 to IT<sup>70–72</sup>. We find that hV4 is responsive to both disparity and motion cues (although only in Experiment 2), which is consistent with reports that macaque V4 is sensitive to both relative disparity<sup>34,35,73</sup>, and relative motion<sup>74</sup>. In the macaque ventral stream, areas down-stream of V4 include sub-divisions of IT, TEO and TE, that are responsive to relative disparity<sup>75,76</sup> and to relative motion<sup>22,77</sup>. In our data, responses to both motion- and disparity-defined form are weak or unmeasurable in ventral areas downstream of hV4 (VO1, VO2, PHC1, PHC2), but strong in areas that lie on the lateral surface (LO1, LO2 and V3B). This difference in responsiveness is not likely due to instrumental factors as the contrast-based localizer condition produces robust responses in at least three of the ventral areas (VO1, VO2 and PHC1). Thus, for motion- and disparity-defined forms, the functional homology is poor between human ventral-surface areas and the canonical macaque ventral areas, but better between human lateral areas and macaque ventral areas. Two recent proposals divide human visual cortex into three streams (dorsal, ventral and lateral), rather than the two canonical dorsal and ventral streams of the macaque<sup>78,79</sup>. Our data suggests that a functional homology exists between human lateral areas and the macaque ventral areas in IT cortex. Human ventral surface areas may either lack clear homologues in the macaque, or the homologous areas remain to be discovered.

## Methods

**Participants.** 15 healthy adult participants (5 female; mean age = 30.6 ± 13.5) participated in Experiment 1 and 14 participated in Experiment 2 (7 female; mean age = 31.9 ± 14.2), with 8 participants taking part in both experiments. Each participant had visual acuity better than +0.1 LogMar (20/25) in each eye as measured on a Bailey-Lovie chart and stereo-acuity of 40 seconds of arc or better on the RandDot stereoacuity test. The experiment began after the procedures of the study had been explained and the participant had given written informed consent. Experiment protocol and consent forms were approved by the Stanford University Institutional Review Board, and all methods were performed in accordance with the relevant guidelines and regulations.

**Visual stimuli.** The stimuli for both experiments were shown on a 47” Resonance Technology LCD display and viewed through a mirror at a distance of 277 cm. This resulted in a presentation area of 12.1 × 21.2°/visual angle, of which our stimuli occupied 12 × 12°. The screen resolution was 1024 × 768 pixels, 8-bit color depth and a refresh rate of 60 Hz. In the relative disparity condition, the mean luminance was 2.17 cd/m<sup>2</sup> and contrast was 60%, and stereoscopic stimuli were displayed using red/blue anaglyph glasses, which were worn throughout the experiment. In the other two conditions, the mean luminance was 34.49 cd/m<sup>2</sup> and contrast was 90%.

For each stimulus condition, the display comprised a central 2° radius disk region and an immediately adjacent 6° radius annulus. In the relative motion condition, the disk-boundary was defined using a random-dot

kinematogram. In the relative disparity condition, the disk and the annulus were defined using dynamic random-dot stereograms with no monocular cues. For the relative disparity display, dot size was 5 minutes of arc (arcmin) and dot density was 36 per  $(^\circ/\text{visual angle})^2$ , while in the relative motion display, dot size was 10.4 arcmin and dot density was 10 per  $(^\circ/\text{visual angle})^2$ . In Experiment 1, we also ran a boundary localizer condition in which the disk-annulus boundary was defined by a contrast difference in texture patterns comprised of 1-dimensional noise which alternated between horizontal and vertical orientations at 3 Hz. This condition allowed us to compare the boundary activations found in the motion and disparity conditions to the activations generated by a contrast-defined boundary to which all visual areas should be highly sensitive. This localizer also served to verify the accuracy of the retinotopy template<sup>52</sup> we used.

In the contrast condition, the central disk was presented in what we will refer to as the “A block” of the fMRI design and alternated with the adjacent annulus configuration, presented during the “B block” (see Fig. 1A). In the relative motion condition, the horizontal positions of individual dots comprising a random dot pattern updated at 20 Hz. Dots were displaced by 10 arcmin per update ( $3.33^\circ/\text{sec}$ ) with a dot life-time of 100 video frames. The dots moved leftwards or rightwards, changing direction at 1 Hz. When these direction changes are quadrature phase shifted between the inside and the outside of the central disk, the dots move in opposite directions during half of each cycle, and in the same direction during the other half. When changes occur in phase, dots move in the same direction throughout the cycle. Motion in opposite directions produces a relative motion defined spatially segmented percept with a visible boundary between the disk and annulus regions. Motion in the same direction produces uniform motion with no boundary. In the A block, direction changes were made to occur out of phase between the center and surround, producing square-wave alternations between uniform motion and segmented configurations at 1 Hz. In the B block, direction changes were made to occur in phase. Locally, each part of the display contained dots that alternated between leftward and rightward motion at 1 Hz, only the relative direction of motion over the disk and annulus regions differed between A and B blocks.

In the relative disparity condition, the positions of individual dots updated at 20 Hz such that the dot fields were binocularly correlated, but temporally uncorrelated (no monocular cues). The horizontal disparity of the central disk and the annulus alternated at 2 Hz between 0 disparity and 20 arcmin of uncrossed disparity. In the A block, the disk and annulus alternated in anti-phase, generating a spatially segmented percept with a visible boundary between the disk and annulus regions defined by relative disparity. In the B block, the disk and annulus alternated in phase, leading to a uniform motion percept with no border. Thus, disparity modulated between 0 and 20 arcmin at all locations in both A and B blocks, with only the relative disparity over the disk and annulus regions differing between A and B blocks. Participants wore anaglyph glasses throughout the experiment, but the contrast and motion condition were identical in both eyes and thus effectively shown at 0 disparity.

In Experiment 2 we replicated the relative motion and disparity conditions from Experiment 1, but introduced a rapid serial visual presentation (RSVP) task at fixation that served to direct attention away from the stimulus. Subjects attended to an array of five letters, presented in trials of three array updates. Each trial consisted of a pre-target array of five randomly oriented Fs, followed by a target array, followed by a masker array identical to the pre-target array. The target array consisted of five Ls in the same positions and orientations as the pre-target Fs, but on a fraction of trials a randomly selected L was replaced by a T. On each change, subjects had to indicate with a button-press whether the target array was a T among Ls, or all Ls. The array duration was adapted online using a staircase procedure to stabilize performance at a constant level ( $\sim 80\%$  correct) during both A and B blocks. The letter display was superimposed on the motion and disparity stimuli, covering the central part of the disk without overlapping with the boundary.

**fMRI experimental procedure.** We used a block design in which 12 s A blocks alternated periodically with 12 s B blocks, yielding a 24 s base period for the paradigm that was repeated 10 times in what we refer to as a “scan”. The design is illustrated schematically in Fig. 1. Ten stimulus cycles were shown per scan, with an additional half-cycle (one 12 s control block) being shown in the beginning of the scan to allow the brain and the scanner to settle. The data collected during this “dummy” period were removed from the fMRI time series data before the data analysis. The disparity condition was not run for 2 out of 15 participants in Experiment 1 because of technical issues. We acquired 4 scans per condition for each participant in Experiment 1, except 3/15 participants for whom we acquired only 3 scans for the contrast localizer condition. In Experiment 2, we acquired 4 scans of each condition per participant, except one for whom we only acquired 3 scans of each.

**Structural and functional MRI acquisition.** Functional and structural MRI data were collected on a General Electric Discovery 750 (General Electric Healthcare) equipped with a 32-channel head coil (Nova Medical) at the Center for Cognitive and Neurobiological Imaging at Stanford University. For each participant, we acquired two whole-brain T1-weighted structural datasets ( $1.0 \times 1.0 \times 1.0$  mm resolution, TE = 2.5 ms, TR = 6.6 ms, flip angle = 12, FOV =  $256 \times 256$ ) that were used for tissue segmentation and registration with atlas-based ROIS and retinotopy template (see below). In both experiments, a multiplexed EPI sequence<sup>80</sup> was used for the functional data, which allowed for the collection of 60 horizontal slices ( $2.0 \times 2.0 \times 2.0$  mm resolution, TE = 30 ms, TR = 2000 ms, flip angle = 77, FOV =  $220 \times 220$ ), resulting in whole-brain coverage.

**fMRI preprocessing and analysis.** Preprocessing of anatomical and functional data was performed using fMRIPrep 1.0.8<sup>81</sup> (RRID:SCR\_016216), which is based on Nipype 1.0.0<sup>82</sup> (RRID:SCR\_002502). Each T1w (T1-weighted) volume was corrected for intensity non-uniformity (INU) using N4BiasFieldCorrection<sup>83</sup> (ANTs 2.2.0). A T1w-reference map was computed after registration of the T1w images (after INU-correction) using `mri_robust_template`<sup>84</sup> (FreeSurfer 6.0.1; RRID:SCR\_001847). The T1w-reference was then skull-stripped using `antsBrainExtraction.sh` (ANTs 2.2.0), using OASIS30ANTs as target template. Brain surfaces were reconstructed using FreeSurfer’s `recon-all`<sup>85</sup>, and the brain mask estimated previously was refined with a custom variation of

the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter implemented in Mindboggle<sup>86</sup> (RRID:SCR\_002438). Spatial normalization to the ICBM 152 Nonlinear Asymmetrical template version 2009c<sup>87</sup> (RRID:SCR\_008796) was performed through nonlinear registration with antsRegistration<sup>88</sup> (ANTs 2.2.0, RRID:SCR\_004757), using brain-extracted versions of both T1w volume and template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast<sup>89</sup> (FSL 5.0.9, RRID:SCR\_002823).

The following preprocessing was performed on the functional data. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. A deformation field to correct for susceptibility distortions was estimated based on a field map that was co-registered to the BOLD reference, using custom fMRIPrep workflow derived from D. Greve's epidewarp.fsl script and further improvements of HCP pipelines<sup>90</sup>. Based on the estimated susceptibility distortion, an unwarped BOLD reference was calculated for a more accurate co-registration with the anatomical reference. The BOLD reference was then co-registered to the T1w reference using flirt<sup>91</sup> (FSL 5.0.9) with the boundary-based registration<sup>92</sup> cost-function. Co-registration was configured with nine degrees of freedom to account for distortions remaining in the BOLD reference. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using mcflirt<sup>93</sup> (FSL 5.0.9). The BOLD time-series were then slice time corrected the data using AFNI's 3dTShift<sup>94</sup> such that all slices were realigned in time to the middle of each TR and resampled into their original, native space by applying a single, composite transform that corrected for head-motion and susceptibility distortions. These resampled BOLD time-series were thus in registration with the T1w reference and will be referred to as preprocessed BOLD in original space, or just preprocessed BOLD. Many internal operations of fMRIPrep use Nilearn 0.4.0<sup>95</sup> (RRID:SCR\_001362), mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in fMRIPrep's documentation.

We then used several AFNI<sup>94</sup> functions to perform additional analysis steps: After eliminating dummy TRs, preprocessed BOLD time-series were scaled (each voxel was scaled to a mean of 100, and values were clipped at 200), and de-trended (removing components corresponding to the six head motion parameters, as well as Legendre polynomials of order 0 (constant signal), 1 (linear drift) and 2). Brain surfaces output by recon-all were converted to AFNI SUMA format using the @SUMA\_Make\_Spec\_FS function. Finally, the 3dVol2Surf function was used to project the volume data onto each subject's native surface as well as onto a standardized surface based on an icosahedron with 198812 nodes (linear depth of 141, see<sup>96</sup>). The native surface was used for ROI analysis, while the standardized surface was used for whole-brain analysis.

The remainder of the analysis was performed in custom Python code. The time-course data were first averaged across the scans for each condition, and then across the voxels within each visual region-of-interest (ROI). We then applied a Fourier transform to the average time-course for each ROI, omitted DC, multiplied the spectrum by 2 to get the single sided spectrum, and scaled by dividing with the number of samples in the time-course. We selected the complex value at the stimulus frequency (10 cycles per scan) for each participant, within each ROI, and used it for statistical analysis. We also used the average time-course to compute the mean cycle for each ROI. The order of the A and B block was mistakenly reversed for the disparity condition in Experiment 1, such that the B block was shown first. To facilitate easy comparison with the other conditions, we corrected for this by shifting the phase of the complex value by 180°, and adding 6 NaN-valued dummy TRs in the beginning of each run prior to computing the mean cycle. For the whole-brain analysis (see below), we performed the same Fourier analysis on a voxel-by-voxel basis, without averaging across ROIs, which gave us a complex value at the stimulus frequency, for every voxel in each participant.

**Visual regions-of-interest.** Topographically organized visual ROIs were derived from a probabilistic atlas<sup>97</sup>. The atlas was generated by functionally defining 25 topographic ROIs covering 22 visual areas in ~50 individual participants, standardizing each participants' surface using icosahedral tessellation and projection<sup>96</sup>, and then assessing the likelihood, across participants, of any particular vector on the standardized surface belonging to a particular ROI<sup>97</sup>. The atlas was defined using a maximum probability approach, which considers a given vector as part of the set of ROIs if it is more often found within the set, than outside the set, across participants. If this is the case, the vector is then assigned the value of the most likely ROI, and if not, it is considered to be outside the set of ROIs. The maximum probability approach captures much of the overall structure of ROIs defined for individual subjects and generalizes well to novel participants that did not contribute to the atlas generation<sup>97</sup>.

We downloaded the atlas from [http://scholar.princeton.edu/sites/default/files/napl/files/probatlas\\_v4.zip](http://scholar.princeton.edu/sites/default/files/napl/files/probatlas_v4.zip) and converted the ROIs from standardized surface space to native surface space for each of our participants, using nearest-neighbor interpolation. We removed vertices that were more than 1 edge away from the main cluster of each ROI, to ensure that all ROIs consisted exclusively of contiguous vertices. This step eliminated small speckles, while having minimal effect on the overall structure and extent of the ROIs. We excluded four ROIs from our analysis, IPS4 and 5, TO2 and FEF, because of their small size in the probabilistic atlas. We also excluded the dorsal and ventral segments of V1, V2 and V3, which were analyzed using the sub-ROI approach described below. This gave us a total of 15 bilateral ROIs to analyze.

To derive an independent estimate of the response in regions of early visual cortex responding to different eccentricities in the visual field, we used a template developed by Benson and colleagues<sup>52</sup> that accurately predicts the location and retinotopic organization of early visual areas V1-V3, using only the cortical anatomy. After transforming the template data to match the specific cortical topology of each participant, we could now sub-divide the ROIs in early visual cortex, for each participant, by selecting voxels within the dorsal and ventral segments of V1, V2 and V3, defined based on the probabilistic atlas, that were responsive to a given range of eccentricities. We generated 24 overlapping sub-ROIs for each early visual area, centered on radii ranging from 0.25° to 6.0°, separated by 0.25°, and each spanning 0.5°/vis. angle.

For both ROI analyses, we excluded any participant for whom the combined number of vertices across both hemispheres was less than 50, and any ROI where less than 3 participants reached that criterion was excluded from analysis. This cut-off meant that for V2, the sub-ROI centered at 0.25° was not included in the analysis, while for V3, sub-ROIs centered at 0.25 and 0.5° were not included. For most of the remaining sub-ROIs (and all ROIs outside early visual cortex), all participants survived the criterion and were included. At least 10 participants survived per sub-ROI, except for the sub-ROI centered at 0.25° in V1, which in Experiment 2 only included 8 participants.

**Vector-based statistics.** We computed the average phase and amplitude at the stimulus frequency using a vector-based approach, in which the real and imaginary part of the complex value was averaged separately across participants, and then combined so that vector-average amplitude and phase could be computed. The real and imaginary values used to compute the vector-average amplitudes in Experiments 1 and 2 have been plotted for the most significant negatively and positively signed V3 sub-ROIs, as well as for V3B, in Supplementary Figs 1 and 2. Error bars were computed using a geometrical approach, in which a two-dimensional error ellipse is computed, which describes the standard error of the mean response amplitude. The upper and lower error bounds were computed as the longest and shortest vectors from the origin to the error ellipse (a detailed description of this approach can be found in<sup>98</sup>). Statistical tests for significance of responses in individual ROIs were run as one-sample Hotelling's  $t^2$  tests of the null hypothesis that the two-dimensional data set containing the real and imaginary parts of the complex value at the stimulus frequency was equal to [0,0]<sup>99</sup>. Note that this vector-based approach means that both amplitude and phase, and their consistency across participants, contributes to our reported estimates of mean amplitude, error and statistical significance. Following this logic, we used a two-sample Hotelling's  $t^2$  to compare responses across the two experiments. Both one-sample and two-sample tests were run using the ICSNP<sup>100</sup> package in R<sup>101</sup>.

We multiplied mean amplitude with phase sign, to illustrate the phase of the responses. We computed the phase sign in the following way. First, response phase was converted to response delay in seconds following the onset of the A block, adding 1 second to the results to account for the fact that voxel time-series were shifted to the center of each 2-second TR during slice-timing correction. This response delay, indicating the timing of the peak response, was now used to derive the phase sign of the response. Response delays at no less than 2 seconds into a block and no more than 2 seconds into the next, were associated with that block (given positive signs for A blocks, negative signs for B blocks), under the logic that responses peaking during the first 2-sec TR of a block would more likely be associated with the preceding block. As an additional precaution against ambiguity in the signing of responses, ROIs that peaked within one TR of the crossover-point (i.e. during the first four seconds of a block) were given signs, but not considered significant, even if the one-sample test was significant at  $\alpha = 0.05$ . The goal was to ensure that positively and negatively signed ROIs that were considered significant peaked at least 2 TRs apart. In Experiment 1, this procedure affected only LO-2 in the contrast condition, and none of the sub-ROIs. In the disparity condition of Experiment 2, seven V1 sub-ROIs were affected (centered on: 2.5°, 2.75°, 3°, 3.25°, 3.5°, 3.75° and 4°) and one V2 sub-ROI (centered on: 2.75°). The procedure affected no other ROIs or sub-ROIs that otherwise reached significance.

**Whole-brain analysis.** To provide an overview of the effect of our conditions across the whole brain, and account for any potential effects outside our set of visual ROIs, we mapped the complex values at each voxel onto a standardized cortical surface based on a icosahedron with 198812 nodes (linear depth of 141, see<sup>96</sup>), for each participant. This surface-based normalization across participants offers several advantages over volume-based approaches to group analysis<sup>102</sup>, most importantly by considering the structure of cortical sulci and gyri, as opposed to Talairach registration and other types cross-subject normalization in volume-space, which is likely to blur activations across neighboring banks of a sulcus<sup>103</sup>. We then applied the cross-subject analysis that we used for the ROIs and sub-ROIs, at each node of the standardized surface, using the real and imaginary values to compute the vector-average amplitude and response delay, and run the one-sample Hotelling's  $t^2$  test for significance. We then plotted the response delay in seconds for each node on the surface, and thresholded the data at  $\alpha = 0.05$  based on the Hotelling's test. We also applied the same restriction against ambiguous responses as described for the ROIs, and did not plot nodes that peaked during the first four seconds of a block.

## Data Availability

The raw data from Experiments 1 and 2 are available at <http://openneuro.org/datasets/ds001984> and <http://openneuro.org/datasets/ds001978>, respectively.

## References

- Poort, J. *et al.* The Role of Attention in Figure-Ground Segregation in Areas V1 and V4 of the Visual Cortex. *Neuron* **75**, 143–156, <https://doi.org/10.1016/j.neuron.2012.04.032> (2012).
- Poort, J., Self, M. W., van Vugt, B., Malkki, H. & Roelfsema, P. R. Texture Segregation Causes Early Figure Enhancement and Later Ground Suppression in Areas V1 and V4 of Visual Cortex. *Cerebral cortex* **26**, 3964–3976, <https://doi.org/10.1093/cercor/bhw235> (2016).
- Schnabel, U. H. *et al.* Figure-ground perception in the awake mouse and neuronal activity elicited by figure-ground stimuli in primary visual cortex. *Scientific reports* **8**, 17800, <https://doi.org/10.1038/s41598-018-36087-8> (2018).
- Thielscher, A., Kölle, M., Neumann, H., Spitzer, M. & Grön, G. Texture segmentation in human perception: A combined modeling and fMRI study. *Neuroscience* **151**, 730–736, <https://doi.org/10.1016/j.neuroscience.2007.11.040> (2008).
- Scholte, H. S., Jolij, J., Fahrenfort, J. J. & Lamme, V. A. F. Feedforward and Recurrent Processing in Scene Segmentation: Electroencephalography and Functional Magnetic Resonance Imaging. *Journal of cognitive neuroscience* **20**, 2097–2109, <https://doi.org/10.1162/jocn.2008.20142> (2008).

6. Roelfsema, P. R. Cortical Algorithms For Perceptual Grouping. *Annual review of neuroscience* **29**, 203–227, <https://doi.org/10.1146/annurev.neuro.29.051605.112939> (2006).
7. Lamme, V. The neurophysiology of figure-ground segregation in primary visual cortex. *The Journal of Neuroscience* **15**, 1605–1615, <https://doi.org/10.1523/jneurosci.15-02-01605.1995> (1995).
8. Seymour, K. *et al.* Altered Contextual Modulation of Primary Visual Cortex Responses in Schizophrenia. *Neuropsychopharmacology* **38**, 2607, <https://doi.org/10.1038/npp.2013.168> (2013).
9. Ban, H., Preston, T. J., Meeson, A. & Welchman, A. E. The integration of motion and disparity cues to depth in dorsal visual cortex. *Nature neuroscience* **15**, 636–643, <http://www.nature.com/neuro/journal/v15/n4/abs/nn.3046.html> - supplementary-information (2012).
10. Bradshaw, M. F. & Rogers, B. J. The interaction of binocular disparity and motion parallax in the computation of depth. *Vision Res* **36**, 3457–3468 (1996).
11. Nawrot, M. & Blake, R. Neural integration of information specifying structure from stereopsis and motion. *Science* **244**, 716–718, <https://doi.org/10.1126/science.2717948> (1989).
12. Lamme, V. A. F., van Dijk, B. W. & Spekreijse, H. Contour from motion processing occurs in primary visual cortex. *Nature* **363**, 541–543 (1993).
13. Cao, A. N. & Schiller, P. H. Neural responses to relative speed in the primary visual cortex of rhesus monkey. *Visual neuroscience* **20**, 77–84, <https://doi.org/10.1017/S0952523803201085> (2003).
14. Lui, L. L., Bourne, J. A. & Rosa, M. G. P. Single-unit responses to kinetic stimuli in New World monkey area V2: Physiological characteristics of cue-invariant neurones. *Experimental Brain Research* **162**, 100–108, <https://doi.org/10.1007/s00221-004-2113-9> (2005).
15. Shen, Z.-M., Xu, W.-F. & Li, C.-Y. Cue-invariant detection of centre-surround discontinuity by V1 neurons in awake macaque monkey. *The Journal of Physiology* **583**, 581–592, <https://doi.org/10.1113/jphysiol.2007.130294> (2007).
16. Yin, J. *et al.* Breaking cover: neural responses to slow and fast camouflage-breaking motion. *Proceedings of the Royal Society B: Biological Sciences* **282**, <https://doi.org/10.1098/rspb.2015.1182> (2015).
17. Chen, M. *et al.* An Orientation Map for Motion Boundaries in Macaque V2. *Cerebral cortex* **26**, 279–287, <https://doi.org/10.1093/cercor/bhu235> (2016).
18. Allman, J., Miezin, F. & McGuinness, E. Direction- and Velocity-Specific Responses from beyond the Classical Receptive Field in the Middle Temporal Visual Area (MT). *Perception* **14**, 105–126, <https://doi.org/10.1068/p140105> (1985).
19. Marcar, V. L., Xiao, D. K., Raiguel, S. E., Maes, H. & Orban, G. A. Processing of kinetically defined boundaries in the cortical motion area MT of the macaque monkey. *J Neurophysiol* **74**, 1258–1270 (1995).
20. Raiguel, S., Van Hulle, M. M., Xiao, D. K., Marcar, V. L. & Orban, G. A. Shape and spatial distribution of receptive fields and antagonistic motion surrounds in the middle temporal area (V5) of the macaque. *Eur J Neurosci* **7**, 2064–2082 (1995).
21. Xiao, D. K., Raiguel, S., Marcar, V. & Orban, G. A. The spatial distribution of the antagonistic surround of MT/V5 neurons. *Cerebral Cortex* **7**, 662–677 (1997).
22. Sary, G., Vogels, R., Kovacs, G. & Orban, G. A. Responses of monkey inferior temporal neurons to luminance-, motion-, and texture-defined gratings. *Journal of neurophysiology* **73**, 1341–1354 (1995).
23. Marcar, V. L., Raiguel, S. E., Xiao, D. & Orban, G. A. Processing of Kinetically Defined Boundaries in Areas V1 and V2 of the Macaque Monkey. *Journal of neurophysiology* **84**, 2786–2798 (2000).
24. Dupont, P. *et al.* The kinetic occipital region in human visual cortex. *Cerebral cortex* **7**, 283–292 (1997).
25. Van Oostende, S., Sunaert, S., Van Hecke, P., Marchal, G. & Orban, G. A. The kinetic occipital (KO) region in man: an fMRI study. *Cerebral cortex* **7**, 690–701 (1997).
26. Reppas, J. B., Niyogi, S., Dale, A. M., Sereno, M. I. & Tootell, R. B. H. Representation of motion boundaries in retinotopic human visual cortical areas. *Nature* **388**, 175–179 (1997).
27. Shulman, G. L., Schwarz, J., Miezin, F. M. & Petersen, S. E. Effect of Motion Contrast on Human Cortical Responses to Moving Stimuli. *Journal of neurophysiology* **79**, 2794–2803 (1998).
28. Skiera, G., Petersen, D., Skalej, M. & Fahle, M. Correlates of figure-ground segregation in fMRI. *Vision research* **40**, 2047–2056, [https://doi.org/10.1016/S0042-6989\(00\)00038-9](https://doi.org/10.1016/S0042-6989(00)00038-9) (2000).
29. Larsson, J., Heeger, D. J. & Landy, M. S. Orientation Selectivity of Motion-Boundary Responses in Human Visual Cortex. *Journal of neurophysiology* **104**, 2940–2950, <https://doi.org/10.1152/jn.00400.2010> (2010).
30. Thomas, O. M., Cumming, B. G. & Parker, A. J. A specialization for relative disparity in V2. *Nature neuroscience* **5**, 472–478, <https://doi.org/10.1038/nn837> (2002).
31. Bredfeldt, C. E. & Cumming, B. G. A simple account of cyclopean edge responses in macaque v2. *J Neurosci* **26**, 7581–7596, <https://doi.org/10.1523/JNEUROSCI.5308-05.2006> (2006).
32. Qiu, F. T. & von der Heydt, R. Figure and ground in the visual cortex: v2 combines stereoscopic cues with gestalt rules. *Neuron* **47**, 155–166, <https://doi.org/10.1016/j.neuron.2005.05.028> (2005).
33. Anzai, A., Chowdhury, S. A. & DeAngelis, G. C. Coding of Stereoscopic Depth Information in Visual Areas V3 and V3A. *The Journal of Neuroscience* **31**, 10270–10282, <https://doi.org/10.1523/jneurosci.5956-10.2011> (2011).
34. Umeda, K., Tanabe, S. & Fujita, I. Representation of Stereoscopic Depth Based on Relative Disparity in Macaque Area V4. *Journal of neurophysiology* **98**, 241–252, <https://doi.org/10.1152/jn.01336.2006> (2007).
35. Shiozaki, H. M., Tanabe, S., Doi, T. & Fujita, I. Neural Activity in Cortical Area V4 Underlies Fine Disparity Discrimination. *The Journal of Neuroscience* **32**, 3830–3841, <https://doi.org/10.1523/jneurosci.5083-11.2012> (2012).
36. Janssen, P., Vogels, R., Liu, Y. & Orban, G. A. Macaque Inferior Temporal Neurons Are Selective for Three-Dimensional Boundaries and Surfaces. *The Journal of Neuroscience* **21**, 9419–9429 (2001).
37. Krug, K. & Parker, A. J. Neurons in Dorsal Visual Area V5/MT Signal Relative Disparity. *The Journal of Neuroscience* **31**, 17892–17904, <https://doi.org/10.1523/jneurosci.2658-11.2011> (2011).
38. Tsao, D. Y. *et al.* Stereopsis Activates V3A and Caudal Intraparietal Areas in Macaques and Humans. *Neuron* **39**, 555–568, [https://doi.org/10.1016/S0896-6273\(03\)00459-8](https://doi.org/10.1016/S0896-6273(03)00459-8) (2003).
39. Backus, B. T., Fleet, D. J., Parker, A. J. & Heeger, D. J. Human Cortical Activity Correlates With Stereoscopic Depth Perception. *Journal of neurophysiology* **86**, 2054–2068 (2001).
40. Mendola, J. D., Dale, A. M., Fischl, B., Liu, A. K. & Tootell, R. B. H. The Representation of Illusory and Real Contours in Human Cortical Visual Areas Revealed by Functional Magnetic Resonance Imaging. *The Journal of Neuroscience* **19**, 8560–8572 (1999).
41. Minini, L., Parker, A. J. & Bridge, H. Neural Modulation by Binocular Disparity Greatest in Human Dorsal Visual Stream. *Journal of neurophysiology* **104**, 169–178, <https://doi.org/10.1152/jn.00790.2009> (2010).
42. Neri, P., Bridge, H. & Heeger, D. J. Stereoscopic Processing of Absolute and Relative Disparity in Human Visual Cortex. *Journal of neurophysiology* **92**, 1880–1891, <https://doi.org/10.1152/jn.01042.2003> (2004).
43. Grossberg, S. Cortical Dynamics of Figure-Ground Separation in Response to 2D Pictures and 3D Scenes: How V2 Combines Border Ownership, Stereoscopic Cues, and Gestalt Grouping Rules. *Front Psychol* **6**, 2054, <https://doi.org/10.3389/fpsyg.2015.02054> (2015).
44. Raudies, F. & Neumann, H. A neural model of the temporal dynamics of figure-ground segregation in motion perception. *Neural networks: the official journal of the International Neural Network Society* **23**, 160–176, <https://doi.org/10.1016/j.neunet.2009.10.005> (2010).

45. Neumann, H., Yazdanbakhsh, A. & Mignolla, E. Seeing surfaces: the brain's vision of the world. *Physics of Life Reviews* **4**, 189–222 (2007).
46. Craft, E., Schutze, H., Niebur, E. & von der Heydt, R. A neural model of figure-ground organization. *J Neurophysiol* **97**, 4310–4326 (2007).
47. Adelson, E. H. & Bergen, J. R. Spatiotemporal energy models for the perception of motion. *J Opt Soc Am A* **2**, 284–299 (1985).
48. Ohzawa, I. Mechanisms of stereoscopic vision: the disparity energy model. *Curr Opin Neurobiol* **8**, 509–515 (1998).
49. Vinberg, J. & Grill-Spector, K. Representation of Shapes, Edges, and Surfaces Across Multiple Cues in the Human Visual Cortex. *Journal of neurophysiology* **99**, 1380–1393, <https://doi.org/10.1152/jn.01223.2007> (2008).
50. Aissani, C., Cottureau, B., Dumas, G., Paradis, A. L. & Lorenceau, J. Magnetoencephalographic signatures of visual form and motion binding. *Brain research* **1408**, 27–40, <https://doi.org/10.1016/j.brainres.2011.05.051> (2011).
51. Cottureau, B. R., McKee, S. P., Ales, J. M. & Norcia, A. M. Disparity-Specific Spatial Interactions: Evidence from EEG Source Imaging. *The Journal of Neuroscience* **32**, 826–840, <https://doi.org/10.1523/jneurosci.2709-11.2012> (2012).
52. Benson, N. C., Butt, O. H., Brainard, D. H. & Aguirre, G. K. Correction of Distortion in Flattened Representations of the Cortical Surface Allows Prediction of V1-V3 Functional Organization from Anatomy. *PLoS computational biology* **10**, e1003538, <https://doi.org/10.1371/journal.pcbi.1003538> (2014).
53. Gandhi, S. P., Heeger, D. J. & Boynton, G. M. Spatial attention affects brain activity in human primary visual cortex. *Proceedings of the National Academy of Sciences* **96**, 3314–3319 (1999).
54. Martinez, A. *et al.* Involvement of striate and extrastriate visual cortical areas in spatial attention. *Nature neuroscience* **2**, 364–369 (1999).
55. Somers, D. C., Dale, A. M., Seiffert, A. E. & Tootell, R. B. Functional MRI reveals spatially specific attentional modulation in human primary visual cortex. *Proceedings of the National Academy of Sciences* **96**, 1663–1668 (1999).
56. Likova, L. T. & Tyler, C. W. Occipital network for figure/ground organization. *Experimental Brain Research* **189**, 257, <https://doi.org/10.1007/s00221-008-1417-6> (2008).
57. Harvey, B. M. & Dumoulin, S. O. The relationship between cortical magnification factor and population receptive field size in human visual cortex: constancies in cortical architecture. *The Journal of neuroscience: the official journal of the Society for Neuroscience* **31**, 13604–13612, <https://doi.org/10.1523/JNEUROSCI.2572-11.2011> (2011).
58. Shmuel, A., Augath, M., Oeltermann, A. & Logothetis, N. K. Negative functional MRI response correlates with decreases in neuronal activity in monkey visual area V1. *Nature neuroscience* **9**, 569, <https://doi.org/10.1038/nn1675>, <https://www.nature.com/articles/nn1675-supplementary-information> (2006).
59. Shmuel, A. *et al.* Sustained Negative BOLD, Blood Flow and Oxygen Consumption Response and Its Coupling to the Positive Response in the Human Brain. *Neuron* **36**, 1195–1210, [https://doi.org/10.1016/S0896-6273\(02\)01061-9](https://doi.org/10.1016/S0896-6273(02)01061-9) (2002).
60. Saad, Z. S., DeYoe, E. A. & Ropella, K. M. Estimation of fMRI response delays. *NeuroImage* **18**, 494–504, [https://doi.org/10.1016/S1053-8119\(02\)00024-1](https://doi.org/10.1016/S1053-8119(02)00024-1) (2003).
61. Dougherty, R. F. *et al.* Visual field representations and locations of visual areas V1/2/3 in human visual cortex. *Journal of vision* **3**, 586–598, <https://doi.org/10.1167/3.10.1> (2003).
62. Bridge, H. & Parker, A. J. Topographical representation of binocular depth in the human visual cortex using fMRI. *Journal of vision* **7**, 15–15, <https://doi.org/10.1167/7.14.15> (2007).
63. Cacciamani, L., Scalf, P. E. & Peterson, M. A. Neural evidence for competition-mediated suppression in the perception of a single object. *Cortex* **72**, 124–139, <https://doi.org/10.1016/j.cortex.2015.05.018> (2015).
64. Lamme, V. A. F., Zipser, K. & Spekreijse, H. Figure-ground activity in primary visual cortex is suppressed by anesthesia. *Proceedings of the National Academy of Sciences* **95**, 3263–3268, <https://doi.org/10.1073/pnas.95.6.3263> (1998).
65. Rockland, K. S. & Van Hoesen, G. W. Direct temporal-occipital feedback connections to striate cortex (V1) in the macaque monkey. *Cerebral cortex* **4**, 300–313 (1994).
66. Bair, W., Cavanaugh, J. R. & Movshon, J. A. Time Course and Time-Distance Relationships for Surround Suppression in Macaque V1 Neurons. *The Journal of Neuroscience* **23**, 7690–7701, <https://doi.org/10.1523/jneurosci.23-20-07690.2003> (2003).
67. Breidfeldt, C. E. & Cumming, B. G. A Simple Account of Cyclopean Edge Responses in Macaque V2. *The Journal of Neuroscience* **26**, 7581–7596, <https://doi.org/10.1523/jneurosci.5308-05.2006> (2006).
68. Lamme, V. A. F., Rodriguez-Rodriguez, V. & Spekreijse, H. Separate Processing Dynamics for Texture Elements, Boundaries and Surfaces in Primary Visual Cortex of the Macaque Monkey. *Cerebral cortex* **9**, 406–413, <https://doi.org/10.1093/cercor/9.4.406> (1999).
69. Bakin, J. S., Nakayama, K. & Gilbert, C. D. Visual responses in monkey areas V1 and V2 to three-dimensional surface configurations. *The Journal of neuroscience: the official journal of the Society for Neuroscience* **20**, 8188–8198 (2000).
70. Neri, P. A Stereoscopic Look at Visual Cortex. *Journal of neurophysiology* **93**, 1823–1826, <https://doi.org/10.1152/jn.01068.2004> (2005).
71. Parker, A. J. Binocular depth perception and the cerebral cortex. *Nature Reviews Neuroscience* **8**, 379, <https://doi.org/10.1038/nrn2131> (2007).
72. Verhoef, B.-E., Vogels, R. & Janssen, P. Binocular depth processing in the ventral visual pathway. *Philosophical Transactions of the Royal Society B: Biological Sciences* **371**, 20150259, <https://doi.org/10.1098/rstb.2015.0259> (2016).
73. Tanabe, S., Umeda, K. & Fujita, I. Rejection of false matches for binocular correspondence in macaque visual cortical area V4. *The Journal of neuroscience* **24**, 8170–8180 (2004).
74. Mysore, S. G., Vogels, R., Raiguel, S. E. & Orban, G. A. Processing of Kinetic Boundaries in Macaque V4. *Journal of neurophysiology* **95**, 1864–1880, <https://doi.org/10.1152/jn.00627.2005> (2006).
75. Verhoef, B.-E., Vogels, R. & Janssen, P. Inferotemporal Cortex Suberves Three-Dimensional Structure Categorization. *Neuron* **73**, 171–182, <https://doi.org/10.1016/j.neuron.2011.10.031> (2012).
76. Uka, T., Tanabe, S., Watanabe, M. & Fujita, I. Neural Correlates of Fine Depth Discrimination in Monkey Inferior Temporal Cortex. *The Journal of Neuroscience* **25**, 10796–10802, <https://doi.org/10.1523/jneurosci.1637-05.2005> (2005).
77. Unno, S., Handa, T., Nagasaka, Y., Inoue, M. & Mikami, A. Modulation of neuronal activity with cue-invariant shape discrimination in the primate superior temporal sulcus. *Neuroscience* **268**, 221–235, <https://doi.org/10.1016/j.neuroscience.2014.03.024> (2014).
78. Weiner, K. S. & Grill-Spector, K. Neural representations of faces and limbs neighbor in human high-level visual cortex: evidence for a new organization principle. *Psychological research* **77**, 74–97, <https://doi.org/10.1007/s00426-011-0392-x> (2013).
79. Haak, K. V. & Beckmann, C. F. Objective analysis of the topological organization of the human cortical visual connectome suggests three visual pathways. *Cortex* **98**, 73–83, <https://doi.org/10.1016/j.cortex.2017.03.020> (2018).
80. Feinberg, D. A. *et al.* Multiplexed echo planar imaging for sub-second whole brain fMRI and fast diffusion imaging. *PLoS one* **5**, e15710, <https://doi.org/10.1371/journal.pone.0015710> (2010).
81. Esteban, O. *et al.* fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nature Methods* **16**, 111–116, <https://doi.org/10.1038/s41592-018-0235-4> (2019).
82. Gorgolewski, K. *et al.* Nipype: A Flexible, Lightweight and Extensible Neuroimaging Data Processing Framework in Python. *Frontiers in neuroinformatics* **5**, <https://doi.org/10.3389/fninf.2011.00013> (2011).
83. Tustison, N. J. *et al.* N4ITK: Improved N3 Bias Correction. *IEEE Transactions on Medical Imaging* **29**, 1310–1320, <https://doi.org/10.1109/TMI.2010.2046908> (2010).

84. Reuter, M., Rosas, H. D. & Fischl, B. Highly accurate inverse consistent registration: A robust approach. *NeuroImage* **53**, 1181–1196, <https://doi.org/10.1016/j.neuroimage.2010.07.020> (2010).
85. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage* **9**, 179–194 (1999).
86. Klein, A. *et al.* Mindboggling morphometry of human brains. *PLoS computational biology* **13**, e1005350, <https://doi.org/10.1371/journal.pcbi.1005350> (2017).
87. Fonov, V. S., Evans, A. C., McKinstry, R. C., Almlí, C. R. & Collins, D. L. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage* **47**, S102, [https://doi.org/10.1016/S1053-8119\(09\)70884-5](https://doi.org/10.1016/S1053-8119(09)70884-5) (2009).
88. Avants, B. B., Epstein, C. L., Grossman, M. & Gee, J. C. Symmetric diffeomorphic image registration with cross-correlation: Evaluating automated labeling of elderly and neurodegenerative brain. *Medical Image Analysis* **12**, 26–41, <https://doi.org/10.1016/j.media.2007.06.004> (2008).
89. Zhang, Y., Brady, M. & Smith, S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE Transactions on Medical Imaging* **20**, 45–57, <https://doi.org/10.1109/42.906424> (2001).
90. Glasser, M. F. *et al.* The minimal preprocessing pipelines for the Human Connectome Project. *NeuroImage* **80**, 105–124, <https://doi.org/10.1016/j.neuroimage.2013.04.127> (2013).
91. Jenkinson, M. & Smith, S. A global optimisation method for robust affine registration of brain images. *Medical Image Analysis* **5**, 143–156, [https://doi.org/10.1016/S1361-8415\(01\)00036-6](https://doi.org/10.1016/S1361-8415(01)00036-6) (2001).
92. Greve, D. N. & Fischl, B. Accurate and robust brain image alignment using boundary-based registration. *NeuroImage* **48**, 63–72, <https://doi.org/10.1016/j.neuroimage.2009.06.060> (2009).
93. Jenkinson, M., Bannister, P., Brady, M. & Smith, S. Improved Optimization for the Robust and Accurate Linear Registration and Motion Correction of Brain Images. *NeuroImage* **17**, 825–841, <https://doi.org/10.1006/nimg.2002.1132> (2002).
94. Cox, R. W. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical research* **29**, 162–173 (1996).
95. Abraham, A. *et al.* Machine learning for neuroimaging with scikit-learn. *Frontiers in neuroinformatics* **8**, <https://doi.org/10.3389/fninf.2014.00014> (2014).
96. Argall, B. D., Saad, Z. S. & Beauchamp, M. S. Simplified intersubject averaging on the cortical surface using SUMA. *Human brain mapping* **27**, 14–27, <https://doi.org/10.1002/hbm.20158> (2006).
97. Wang, H. X., Merriam, E. P., Freeman, J. & Heeger, D. J. Motion Direction Biases and Decoding in Human Visual Cortex. *The Journal of Neuroscience* **34**, 12601–12615, <https://doi.org/10.1523/jneurosci.1034-14.2014> (2014).
98. Pei, F., Baldassi, S., Tsai, J. J., Gerhard, H. E. & Norcia, A. M. Development of contrast normalization mechanisms during childhood and adolescence. *Vision research* **133**, 12–20, <https://doi.org/10.1016/j.visres.2016.03.010> (2017).
99. Anderson, T. W. *An introduction to multivariate statistical analysis*. (Wiley, 1984).
100. Nordhausen, K., Sirkia, S., Oja, H. & Tyler, D. E. (2018).
101. R Core Team. (Vienna, Austria, 2014).
102. Hagler, D. J., Saygin, A. P. & Sereno, M. I. Smoothing and cluster thresholding for cortical surface-based group analysis of fMRI data. *NeuroImage* **33**, 1093–1103, <https://doi.org/10.1016/j.neuroimage.2006.07.036> (2006).
103. Fischl, B., Sereno, M. I. & Dale, A. M. Cortical surface-based analysis: II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage* **9**, 195–207 (1999).

## Acknowledgements

This research was supported by NIH grant EY018775.

## Author Contributions

P.J.K., B.R.C. and A.M.N. designed the study. P.J.K. and B.R.C. collected the data. P.J.K. analyzed the data. P.J.K., B.R.C. and A.M.N. wrote the paper.

## Additional Information

**Supplementary information** accompanies this paper at <https://doi.org/10.1038/s41598-019-45036-y>.

**Competing Interests:** The authors declare no competing interests.

**Publisher's note:** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2019