

# Responses in early visual areas to contour integration are context dependent

Cheng Qiu

Department of Psychology, University of Minnesota,  
Minneapolis, MN, USA



Philip C. Burton

Office of the Associate Dean for Research,  
College of Liberal Arts, University of Minnesota,  
Minneapolis, MN, USA



Daniel Kersten

Department of Psychology, University of Minnesota,  
Minneapolis, MN, USA



Cheryl A. Olman

Department of Psychology, University of Minnesota,  
Minneapolis, MN, USA



It has been shown that early visual areas are involved in contour processing. However, it is not clear how local and global context interact to influence responses in those areas, nor has the interarea coordination that yields coherent structural percepts been fully studied, especially in human observers. In this study, we used functional magnetic resonance imaging (fMRI) to measure activity in early visual cortex while observers performed a contour detection task in which alignment of Gabor elements and background clutter were manipulated. Six regions of interest (two regions, containing either the cortex representing the target or the background clutter, in each of areas V1, V2, and V3) were predefined using separate target versus background functional localizer scans. The first analysis using a general linear model showed that in the presence of background clutter, responses in V1 and V2 target regions of interest were significantly stronger to aligned than unaligned contours, whereas when background clutter was absent, no significant difference was observed. The second analysis using interarea correlations showed that with background clutter, there was an increase in V1–V2 coordination within the target regions when perceiving aligned versus unaligned contours; without clutter, however, correlations between V1 and V2 were similar no matter whether aligned contours were present or not. Both the average response magnitude and the connectivity analysis suggest different mechanisms support contour processing with or without background distractors. Coordination between V1 and V2 may play a major role in coherent

structure perception, especially with complex scene organization.

## Introduction

Contour integration involves grouping local features across several levels of abstraction and a range of spatial scales. Small, similar elements positioned closely along an invisible smooth path are perceptually organized as due to a continuous contour. This grouping process is enhanced if the elements have orientations that align with the path (Field, Hayes, & Hess, 1993; Hess & Field, 1999; Kovács, 1996; Li & Gilbert, 2002). Further, knowledge of the global form of the path contributes to local integration, such as the form closing (Kovács, 1996; Kovács & Julesz, 1993, 1994) and smoothness (Pettet, 1999; Pettet, McKee, & Grzywacz, 1998), and the global knowledge is often necessary to disambiguate competing local groupings in cluttered scenes (Ullman & Sha'ashua, 1988).

It has been reported repeatedly that cortical areas higher in the visual hierarchy such as the lateral occipital complex (LOC) and inferotemporal cortex (IT) show selectivity to coherent contours (Altmann, Bühlhoff, & Kourtzi, 2003; Cardin, Friston, & Zeki, 2010; Dumoulin, Dakin, & Hess, 2008; Dumoulin & Hess, 2006; Kourtzi, Tolia, Altmann, Augath, & Logothetis, 2003; Mendola, Dale, Fischl, Liu, & Tootell, 1999; Murray, Kersten, Olshausen, Schrater, &

Citation: Qiu, C., Burton, P. C., Kersten, D., & Olman, C. A. (2016). Responses in early visual areas to contour integration are context dependent. *Journal of Vision*, 16(8):19, 1–18, doi:10.1167/16.8.19.

doi: 10.1167/16.8.19

Received August 24, 2015; published June 30, 2016

ISSN 1534-7362



Woods, 2002; Tanskanen, Saarinen, Parkkonen, & Hari, 2008), which is also consistent with results of shape or object perception from those areas (Fang, Kersten, & Murray, 2008; Grill-Spector, Kourtzi, & Kanwisher, 2001; Grill-Spector, Kushnir, Edelman, Itzhak, & Malach, 1998; Grill-Spector, Kushnir, Hendler et al., 1998; Gross, Rocha-Miranda, & Bender, 1972; Haxby et al., 2001; Kourtzi & Kanwisher, 2000; Logothetis & Sheinberg, 1996; Tanaka, 1996). However, responses to similar contour stimuli in early visual areas are still controversial. While most neurons in the primary visual cortex (V1) are thought to have small receptive fields, it is known that their responses are modulated by contextual information from outside their receptive fields (Allman, Miezin, & McGuinness, 1985; Fitzpatrick, 2000). However, there are substantial conflicting results on how scene context or global perception affects responses in early visual areas, including V1.

Both enhancement (Altmann et al., 2003; Bauer & Heinze, 2002; Chen et al., 2014; Kapadia, Ito, Gilbert, & Westheimer, 1995; Kourtzi et al., 2003; Li, Piëch, & Gilbert, 2006; McManus, Li, & Gilbert, 2011; Roelfsema, Lamme, & Spekreijse, 2004) and suppression (Cardin et al., 2010; Dumoulin & Hess, 2006; Murray et al., 2002; Murray, Schrater, & Kersten, 2004) of cortical responses to coherent contours relative to scrambled elements have been observed in V1 and/or other early visual areas. Recording from individual V1 neurons in monkeys, Kapadia et al. (1995) and Li et al. (2006) showed that multiple randomly placed and oriented line segments outside the neuron's receptive field would inhibit its response to an optimally oriented line within its receptive field; however, once some of the surround segments were placed collinearly with the central line, the response was then facilitated. Similar results have been demonstrated using functional neuroimaging in both monkeys and human subjects (Altmann et al., 2003; Kourtzi et al., 2003): cortical responses from early visual areas including areas V1, V2, and V3 showed selectivity to coherent patterns of closed contours embedded in a field of randomly oriented segments.

In contrast, Cardin et al. (2010) showed lower activity in areas V1/V2 for collinear patterns than for noncollinear ones. Dumoulin and Hess (2006) also showed weaker activity in early visual areas to a 100% coherence circular pattern but stronger activity to scrambled patterns. In a third study, Murray et al. (2002) used line drawings as stimuli and showed smaller responses in V1 to lines that formed two- and three-dimensional shapes than random lines. These were not the only results showing deactivation in early visual areas to stimulus regularities—early visual areas also respond less to coherent than incoherent motion (Händel, Lutzenberger, Thier, & Haarmeier, 2007;

Harrison, Stephan, Rees, & Friston, 2007; McKeefry, Watson, Frackowiak, Fong, & Zeki, 1997). A similar trend has been observed with coherent versus scrambled natural images using functional imaging in human observers (Grill-Spector, Kushnir, Hendler et al., 1998; Lerner, Hendler, Ben-Bashat, Harel, & Malach, 2001; Paradis et al., 2000).

We are interested in how the enhancement and suppression of cortical responses in early visual areas to coherent contours could both be true. In either case, neurons in V1 are responding to the same coherent circular contour, but they show different response patterns based on different studies listed above. We think the response pattern may highly depend on stimulus context: results showing response increase mainly used contours embedded in clutter (i.e., with randomly orientated and placed segments in the background), but the ones showing a decrease were usually using isolated structure or if any, with just uniform background. To test this, we designed a  $2 \times 2$  experiment in which context and contour coherence were manipulated, and we used fMRI to record the blood oxygenation level-dependent (BOLD) signal from the retinotopically corresponding regions in early visual areas (V1, V2, and V3). Observers performed a contour detection task in the experiment. When there was no background clutter present, the BOLD responses in regions corresponding to the location of the target were slightly suppressed by aligned contours compared with the unaligned; however, with background clutter, the responses were significantly enhanced by contour alignment. By analyzing the effect of context and contour coherence on correlations between responses in early visual areas, we further demonstrated that the coordination between the retinotopically relevant regions in V1 and V2 was dependent upon the experimental conditions. These results suggest the involvement of multiple strategies in contour integration, and they interact according to the context.

## Materials and methods

### Participants

Fifteen observers (mean age: 29 years old; nine men, six women) with normal or corrected-to-normal visual acuity participated in the study. The observers provided informed written consent under an experimental protocol that was in accordance with safety guidelines for MRI research and was approved by the Institutional Review Board at the University of Minnesota.

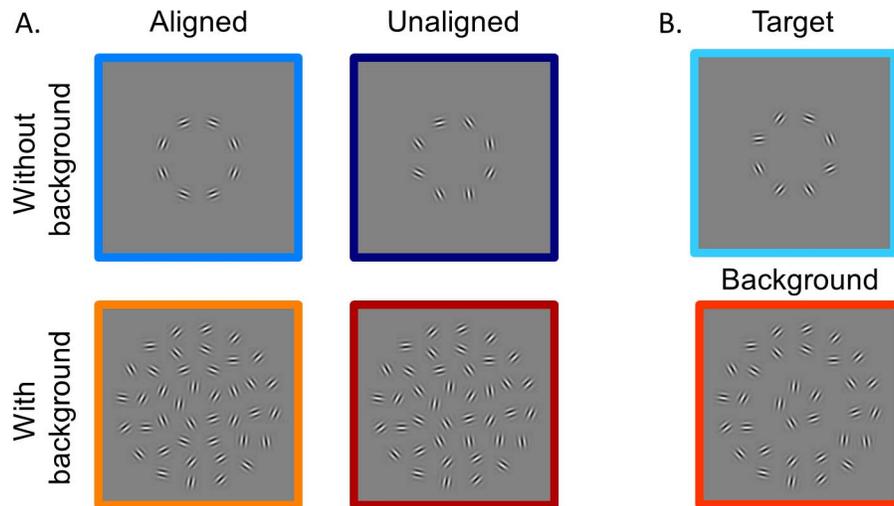


Figure 1. (A) Example stimuli from the four test conditions with a circular contour detection task. (B) Example stimuli from the target (above) versus background (below) differential localizer scan.

## Stimuli

Figure 1 provides examples of the stimuli used, which were generated and presented with Matlab (R2010b; MathWorks, Inc., Natick, MA) using the Psychtoolbox extensions (Brainard, 1997; Pelli, 1997). The target region consisted of eight Gabor patches, which were centered at equal intervals along the circumference of an invisible circle centered at fixation with a radius of  $2^\circ$ ; that is, these patches were evenly spaced at  $2^\circ$  eccentricity from the fixation. Each Gabor patch consisted of a  $4\text{ c}/^\circ$  sinusoidal grating with a random phase offset modulated by a Gaussian envelope with full width at half-maximum of  $0.4^\circ$  ( $\sigma = 0.17^\circ$ ). In an aligned condition, the grating orientation of each Gabor patch was aligned with the tangent line to the invisible circle at the point of patch center. These Gabor patches therefore could be perceived as forming a complete circle. In an unaligned condition, the grating orientation of each Gabor patch was randomly generated. The background clutter consisting of the same Gabor elements as the target was located along invisible circles centered at fixation and with radii of  $1.2^\circ$ ,  $2.9^\circ$ , and  $4.0^\circ$ . Along each circle, the distance between every two Gabor patches was the same as in the target region ( $1.6^\circ$ ), and the numbers of Gabor patches along each eccentricity were 5, 12, and 16, respectively. The grating orientations of these background patches were randomly generated. All Gabor patches had an 80% Michelson contrast and were presented on a mean gray background. The four experimental conditions included aligned target only, unaligned target only, aligned target with background, and unaligned target with background.

## Functional MRI experiments

Stimuli for the first four observers were presented on a NEC 2190UXi monitor with resolution of  $1024 \times 768$  pixels and a refresh rate of 60 Hz. The monitor had a mean luminance of  $110\text{ cd}/\text{m}^2$ . The monitor was mounted to the back wall of the scanning suite; observers viewed the monitor through a mirror mounted to the top of the head coil so that it subtended  $12^\circ$  of visual angle in the horizontal direction and  $9^\circ$  in the vertical direction. For the rest of the observers, stimuli were back-projected via a Sony video projector (spatial resolution of  $1024 \times 768$  pixels, 60 Hz refresh rate, and  $120\text{ cd}/\text{m}^2$  mean luminance) onto a translucent screen placed inside the scanner bore. Observers viewed the stimuli from a distance of 97.5 cm through a mirror located above their eyes (mounted on the head coil), which gave a total image area subtending  $26^\circ \times 20^\circ$ .

Functional MRI data were collected using a 3T Siemens Trio scanner (Siemens, Erlangen, Germany) with a 12-channel head array coil. EPI data were acquired with a field of view  $128\text{ mm} \times 256\text{ mm}$  and a matrix size of  $64 \times 128$  for an in-plane resolution of  $2\text{ mm} \times 2\text{ mm}$ . Slice thickness was 2 mm without interslice gap, and number of slices was 20. Echo time (TE) was 30 ms, repetition time (TR) was 1.5 s, and flip angle was  $80^\circ$ . Four out of the 15 datasets were collected in an axial direction, and the other 11 were in a coronal orientation. Both covered the early visual areas V1, V2, and V3.

A scanning session (Figure 2A) contained three runs (runs 1, 5, and 8) of block-design functional localizers and five event-related runs (runs 2, 3, 4, 6, and 7). Observers were instructed to maintain their fixation on a white square at the center while performing

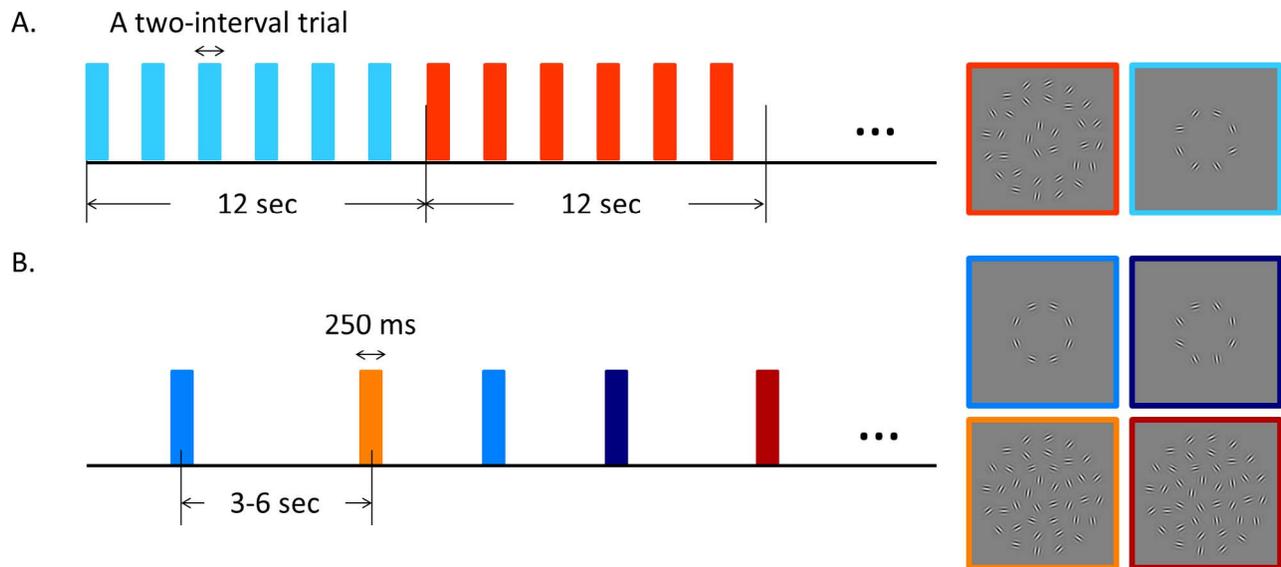


Figure 2. Experimental procedure. (A) An example of block-designed functional localizer scans used to define target or background retinotopically corresponding ROIs. (B) Event-related scans with four experimental conditions.

behavioral tasks during both localizer and event-related scans. Behavioral responses were recorded using a fiber-optic button box (Current Designs, Philadelphia, PA).

Regions of interest (ROIs) were defined by block-designed functional localizer scans (Figure 2A). Each block lasted 12 s, and each run contained 11 “on” blocks alternating with 10 “off” blocks. Each localizer scan thus lasted 252 s. During “on” blocks, randomly oriented Gabor patches at the target region were presented; during “off” blocks only clutter Gabor patches at the background region were presented. The first half-cycle “on” block was discarded before analysis, and the remaining blocks were alternating in 10 cycles per run. During each block, a two-interval trial occurred every 2 s (six trials per block). Duration for both intervals was 200 ms, which were separated by a 200 ms interstimulus interval (ISI). Observers were instructed to press the button when the stimuli from two intervals were the same (with a probability of 12.5%). The fixation square turned green for a correct response and red otherwise.

The event-related runs (Figure 2B) measured BOLD response to four experimental conditions: aligned target only (alnb, aligned/no background), unaligned target only (uanb, unaligned/no background), aligned target with background (albg), and unaligned target with background (uabg). Stimulus duration was 250 ms, and intertrial intervals (ITIs) were 3, 4.5, or 6 s. ITI was randomly assigned and uniformly distributed. Each run included 20 trials for each condition, thus a total of 80 trials per run. The average run length was about 380 s. Observers were required to press the blue button if they perceived an aligned circle, and press the

red button if not. Feedback was provided after each trial.

## Anatomical acquisition and visual area mapping

Prior to the fMRI experiments, each observer participated in a separate retinotopic mapping session, in which a  $T_1$ -weighted anatomical image (MP-RAGE, 1 mm isotropic resolution) was also collected for anatomical reference and cortical surface definition. Gray/white matter segmentation, cortical surface reconstruction, and surface inflation and flattening were completed using FreeSurfer (Dale, Fischl, & Sereno, 1999; Fischl, Sereno, & Dale, 1999). Standard retinotopic mapping including four runs of clockwise/counterclockwise rotating wedges and two runs of expanding/contracting rings (DeYoe et al., 1996; Engel, Glover, & Wandell, 1997; Sereno et al., 1995) was used to identify the early visual areas V1, V2, and V3. Defined visual areas were registered to the reference anatomy for each observer.

## Preprocessing and functional localizers

Functional data was motion corrected using Analysis of Functional NeuroImages software (AFNI; Cox, 1996), with reference to the volume right before a within-session fieldmap image. The motion-corrected data was unwarped using FSL FUGUE to correct distortions introduced by magnetic field inhomogeneities (Smith et al., 2004). High-pass filtering was also applied to the functional localizers data: temporal

frequencies below four cycles per run were removed. The preprocessed functional data were then aligned to anatomical reference data using mrAlign implemented in Matlab (MathWorks, Inc.; <http://gru.stanford.edu/doku.php/mrTools/overview>).

ROIs were defined based on both retinotopic visual areas and functional localizers for each observer. Three repetitions of the functional localizers were averaged to define the ROIs responding to the target (tg) or background (bkgd) regions. For each voxel coherence (unsigned correlation, computed in the Fourier domain as the amplitude of the stimulus-related Fourier component normalized by the square root of the integrated power spectrum) with a sinusoid at the block-alternation frequency, 10 cycles per run, was calculated in the averaged localizer scans (Bandettini, Jesmanowicz, Wong, & Hyde, 1993; Engel et al., 1997). The voxels with coherence exceeding 0.30 were included in the ROIs. The voxels in phase with the target representation were assigned to target ROIs (tgROIs), and the voxels in phase with the background representation were assigned to background ROIs (bkgdROIs). ROIs were initially defined on a flattened cortical surface, where V1, V2, and V3 boundaries could be used to identify the ROIs in different visual areas. Selected voxels were translated to the in-plane space for further refinement to include only contiguous clusters of visually responsive voxels, and the defined ROIs were then exported as a binary mask in the space of the functional data. Six ROIs were defined for each observer: tgV1, tgV2, tgV3, bkgdV1, bkgdV2, and bkgdV3.

## Analysis of the event-related data

### Functional MRI data analysis

Functional image analysis of the event-related runs was conducted using general linear model (GLM) in AFNI with the function 3dDeconvolve. The BOLD response to individual events from each stimulus condition was modeled using the sum of TENT basis functions in 3dDeconvolve, which is a piecewise linear spline function that estimates an impulse response function. The sum of 13 TENT functions was used to cover the duration of 18 s after the stimulus onset, TENT(0, 18, 13). All models were fit separately to each voxel. For each voxel within the predefined ROIs, 13 amplitudes for each stimulus condition were estimated, which were the time course of the estimate hemodynamic response function (HRF) to each condition at the 13 time points (from 0 to 18 s with the time step of 1.5 s). The mean of the first and last two time points was subtracted from each HRF to ensure that it started from and returned to approximately the same baseline level. Estimates from individual voxels were averaged within each of the 6 ROIs, and BOLD response

amplitudes were estimated using the difference between the peak response (reached around 4.5–6 s after stimulus onset) and the baseline response. Our analysis focused on response differences between contour aligned and unaligned conditions, instead of direct comparison between aligned or unaligned contours in clutter versus not, to avoid confound of blood stealing from the background stimulus. Response differences between conditions were assessed using a bootstrapping procedure—resampling the 15 subjects' data with replacement within conditions and calculating differences across 10,000 iterations—and a two-sided permutation test was performed to acquire  $p$  values (Efron & Tibshirani, 1993). Analysis of variance (ANOVA) was also conducted to compare BOLD response amplitudes among conditions in each predefined ROI, and the ANOVA test assumptions of normality and equal variance were checked using the Lilliefors test (Lilliefors, 1967) and Bartlett's test (Snedecor & Cochran, 1956), respectively.

### Connectivity analysis

In addition to the estimated response to each stimulus condition, we also wanted to know whether interregional connections among these ROIs depend on the experimental conditions. Two functional connectivity analysis methods were used to answer this question. The first method, psychophysiological interactions (PPI), uses interaction terms created from the dot product of the seed region time series and vectors representing different experimental conditions to explain variance in time series from other cortical regions with a GLM analysis, which is generalized PPI or gPPI, as in McLaren, Ries, Xu, and Johnson (2012) and Cisler, Bush, and Steele (2014). If the estimated beta weights for the interaction regressors are different among experimental conditions, the coordination between the test ROI and the seed ROI depends on the conditions. In the second method, beta series correlations, a beta weight for each experimental trial is estimated using GLM, these beta weights are sorted according to the experimental condition during the trial, and then correlation coefficients are calculated between the beta weights for each condition (Rissman, Gazzaley, & D'Esposito, 2004). If these correlation coefficients vary depending on the conditions, connectivity between the two regions is condition-dependent. *Psychophysiological interactions:* First, the stimulus effects modeled by the GLM in the previous analysis were subtracted from the preprocessed event-related dataset to generate a residual dataset for further analysis of the intrinsic interactions between cortical areas. Figure 3 shows an example of the PPI terms, which are used as regressors in a GLM. In order to build the psychophysiological interaction terms, both

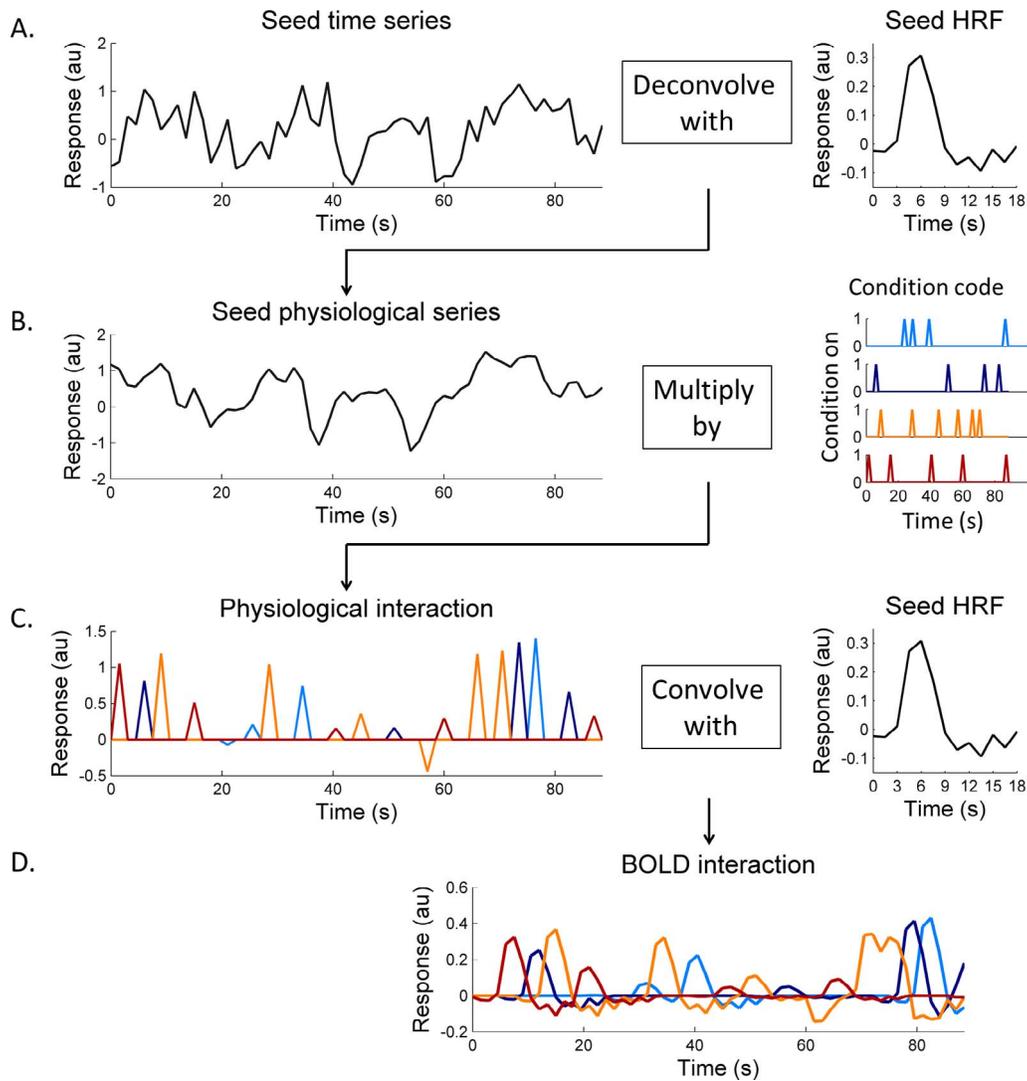


Figure 3. An example of the psychophysiological interactions term. (A) The seed time series were first deconvolved based on the estimated HRF to obtain physiological responses. (B) To combine both the physiological and psychophysical effects we multiplied the seed physiological series by the condition code from each condition separately. (C) The interaction at the physiological level was convolved with the estimated HRF, so we could compare this BOLD level interaction (as shown in D) with residual time series from other ROIs.

psychological condition codes (indicating the representation time for a certain stimulus condition) and physiological responses from the seed region were required. According to McLaren et al. (2012) and Ciser et al. (2014), with more than two conditions, separately building one interaction term for each stimulus condition (gPPI) could be more robust to noise and have better model fits than directly using PPI terms with condition contrast (i.e., “1” for one condition while “-1” for a different condition). Therefore, we created one text file matching the length of functional time series for each condition as the condition code ( $D_c$ ), in which “1” indicated stimulus presentation of that particular condition and “0” for the rest time points (Figure 3B, right panel). The

physiological response from the seed region was estimated by deconvolving the time series from the seed region with its HRF estimated from the previous analysis (Figure 3A). Gitelman, Penny, Ashburner, and Friston (2003) and Kim and Horwitz (2008) demonstrated the importance of modeling the underlying neuronal activity: interaction should be expressed at a neuronal level rather than at the level of hemodynamic responses, with neuronal activity being filtered with an HRF. Interaction was calculated as the product of condition codes and estimated physiological responses (Figure 3B). In order to compare the interaction with BOLD measurements, we then convolved the interaction at a neuronal level with the estimated HRF to

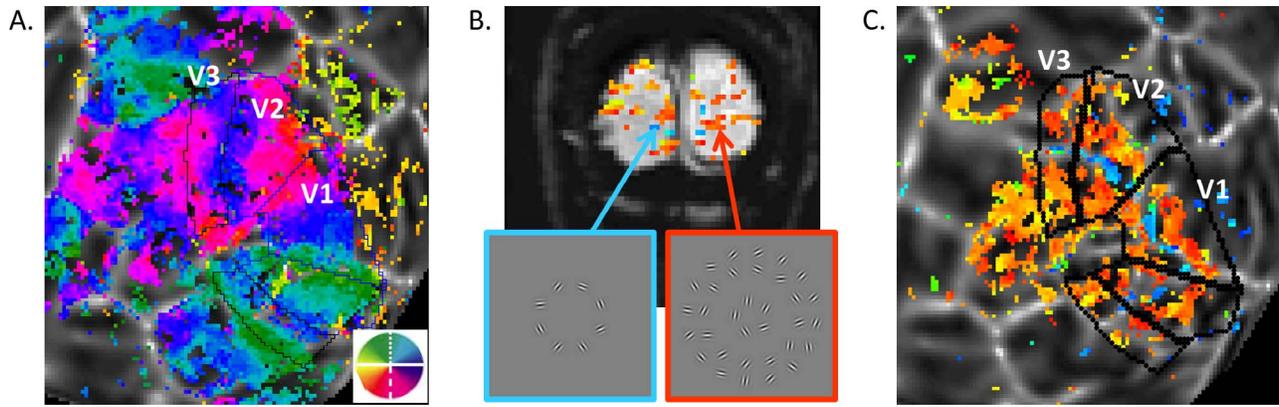


Figure 4. Visual area mapping (A) and functional localizer results (B and C). (A) Angular visual field preference of one observer’s left hemisphere obtained from rotating wedge stimulus (overlay on a flattened patch of the cortical surface centered on the occipital pole). The early visual areas are labeled. (B) On one single coronal EPI image, voxels significantly correlated with the block-alternation are color coded based on relative phases—the bluish voxels are in phase with the target presentation, while the orange voxels are in phase with the background stimulus. (C) Data in B was transformed to the flat patch, where a blue target-associated band and two orange background-associated bands could be seen among the early visual areas.

obtain a regressor at the level of hemodynamic responses (Figure 3C).

In summary, residual time series from the  $i$ th ROI,  $y_i$ , could be modeled as (Friston et al., 1997):

$$y_i = \sum_{c=1}^4 \left( \beta_{c,i} \cdot H(D_c \cdot n_{\text{seed}}) \right) + \beta_{\text{seed},i} \cdot y_{\text{seed}} + \varepsilon_i, \quad (1)$$

where  $c$  is the stimulus condition index, here, we have four conditions,  $c = 1, \dots, 4$ ;  $\beta_{c,i}$  is the correlation coefficient for the PPI regressor of condition  $c$  at ROI  $i$ ;  $\beta_{\text{seed},i}$  is the correlation coefficient for the seed time series regressor at ROI  $i$ ;  $H(\cdot)$  indicates the convolution operation with the estimated TENT HRF for a certain condition ( $c$ ) in the corresponding region (here, the seed region);  $D_c$  is the stimulus presentation code for condition  $c$ ;  $y_{\text{seed}}$  is the time series from the seed region, and  $n_{\text{seed}}$  is the physiological response estimated by deconvolving  $y_{\text{seed}}$  with the estimated HRF; that is, since  $y_{\text{seed}} = H(n_{\text{seed}}) + \text{noise}$ , we could solve for  $n_{\text{seed}}$  given the kernel function and  $y_{\text{seed}}$  using deconvolution. Finally,  $\varepsilon_i$  is an error term at region  $i$ . The beta estimated for the PPI regressor represents the amount of signal that could be explained by both the response in the seed ROI and the stimulus condition. If the beta estimates at a certain ROI from two conditions are different, the seed may differently influence this ROI between these two conditions. Next, 3dDeconvolve and 3dREMLfit (which estimates and removes noise temporal correlations) were used to estimate coefficients based on the model shown in Equation 1. Estimates of  $\beta_{c,i}$  from individual voxels were averaged within ROIs, and would be used for assessing statistical significance.

*Beta series correlations:* With the beta series correlation method (Rissman et al., 2004), we first estimated a beta weight for each experimental trial; that is, modeling the time series from the  $i$ th ROI,  $y_i$ , as

$$y_i = \sum_{c=1}^4 \sum_{j=1}^{c_j} \left( \beta_{i,c,j} \cdot H_i(D_c) \right) + \varepsilon_i, \quad (2)$$

where  $c$  is the index for the stimulus conditions,  $i$  is the index for the ROIs,  $H_i(D_c)$  indicates the estimated hemodynamic response function at the region  $i$  for the condition  $c$ ,  $\beta_{i,c,j}$  is the beta weight for the  $i$ th ROI during the  $j$ th trial of the condition  $c$ ,  $c_j$  is the total number of trials for the condition  $c$ , and  $\varepsilon_i$  is an error term at region  $i$ . Next, the estimated beta values  $\beta_{i,c,j}$  ( $j = 1, \dots, c_j$ ) from the  $i$ th region were regressed against the  $\beta_{i',c,j}$  from the region  $i'$  (the seed) according to the conditions, and the correlation coefficient was used to indicate the connectivity between regions  $i$  and  $i'$  under a certain experimental condition.

## Results

The early visual areas V1, V2, and V3 were manually defined according to the polar angle (Figure 4A) and eccentricity phase maps acquired in separate scanning sessions. Functional localizers (three runs, preprocessed, averaged and subjected to Fourier analysis) were used to define ROIs corresponding to the cortical representations of stimulus target or background regions (Figure 4B, C). On the flat patch, a band of activation associated with the target representation (blue, Figure 4C) and two bands associated with the background representations

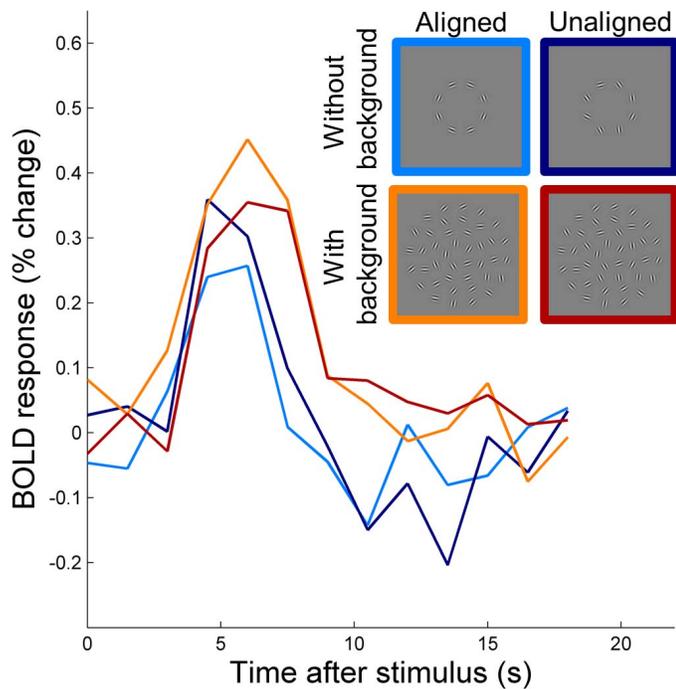


Figure 5. Estimated HRFs from four experimental conditions for one subject’s tgV2 ROI. The HRFs were estimated for each voxel, and then averaged within each ROI.

(orange, Figure 4C) were consistent with the eccentricity features in the early visual areas. Therefore, two sets of ROIs were defined in each visual area corresponding to target (tgROIs) and background regions (bkgdROIs).

### Stimulus-related activity

We first looked at the BOLD response magnitude for each experimental condition. Estimated HRFs in one representative subject within the tgV2 ROI are shown in Figure 5. The differences of BOLD responses to the aligned from the unaligned contours for each ROI depend on the background context as shown in Figure 6. When the background was present, the tgV1 and tgV2 ROIs showed a significant preference for the aligned contours (in tgV1,  $albg - uabg = 0.043$  with two-sided permutation test  $p = 0.012$ , and in tgV2,  $albg - uabg = 0.066$  with two-sided permutation test  $p < 0.001$ ); the bkgdV1 and bkgdV2 ROIs showed a similar trend (in bkgdV1,  $albg - uabg = 0.079$  with two-sided permutation test  $p = 0.028$ , and in bkgdV2,  $albg - uabg = 0.063$  with two-sided permutation test  $p = 0.019$ ). When there was no background, the responses to aligned contours were weaker than the responses to the unaligned contours in tgV1 ( $alnb - uanb = -0.039$ , with two-sided permutation test  $p = 0.13$ ). The tgV3 ROI showed a similar response pattern as in the tgV1 and tgV2 ROIs, but it failed to reach the significance level; this may be related to the simple circular form used in the experiment. Using a two-way ANOVA model (assumptions were satisfied based on the Lilliefors and Bartlett’s test) within each visual area (Alignment and Background as fixed effects, and subjects as a random effect), a significant interaction between Alignment and Background was

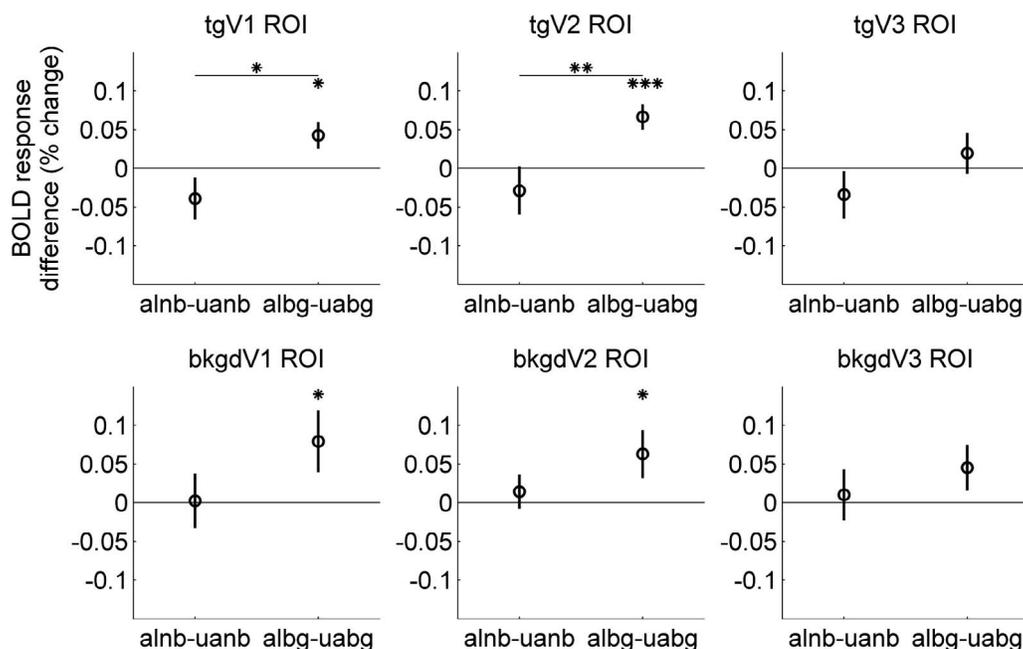


Figure 6. Stimulus-related BOLD response differences among conditions. Each panel shows average data from the 15 observers in one ROI. For each panel, the differences of estimated HRF amplitudes between the aligned and unaligned contours when there was no background are shown on the left, and the differences when there was background clutter are shown on the right. Asterisks indicate statistical significance based on the two-sided permutation test at  $*p < 0.05$ ,  $**p < 0.01$ , and  $***p < 0.001$ . Error bars show  $\pm 1$  SE.

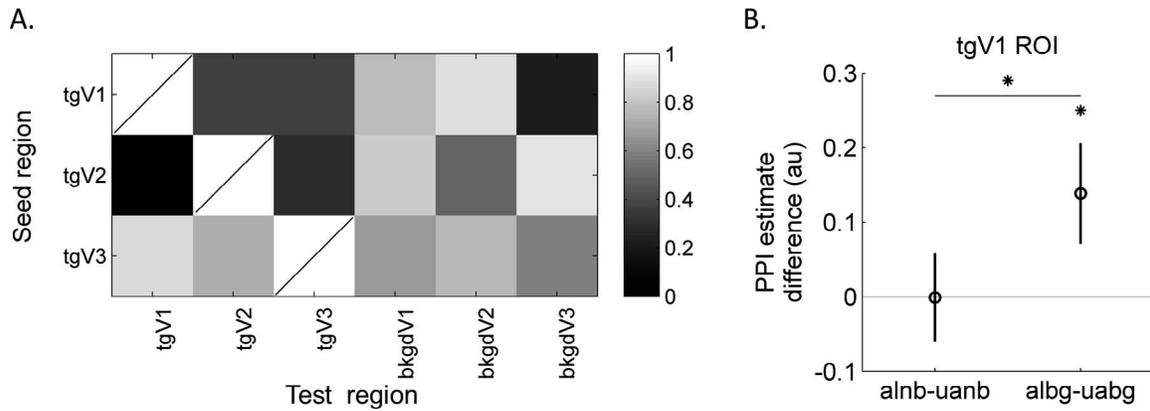


Figure 7. Connectivity results using PPI. (A) The permutation test  $p$  values for interactions of PPI estimates between the Alignment and Background. A stronger interaction (smaller  $p$  value) is shown in darker color. Row indicates the seed region. The ROI is not tested on itself. (B) The differences of Beta estimates among conditions in tgV1 are shown when tgV2 was the seed. The mean difference from the 15 observers of estimated beta weights of PPI terms between the aligned and unaligned contours when there was no background is shown on the left, and the difference when there was background clutter was on the right. Error bars show  $\pm 1$  SE. Asterisks show significant levels based on the permutation test at  $*p < 0.05$ .

observed in tgV2 ROI,  $F(1, 59) = 7.83$ ,  $p = 0.014$ , and a similar trend was seen in tgV1 ROI,  $F(1, 59) = 5.85$ ,  $p = 0.030$ . When we tested for this interaction using a permutation test, tgV2 showed  $p = 0.0062$ , and tgV1 showed  $p = 0.011$ . No significant interaction was found in background ROIs. Thus, the regions in V1 and V2 corresponding retinotopically to the target ring responded to the coherent contour differently according to the context.

### Coordination among regions influenced by physiological and psychological states

We also explored coordination among defined ROIs using two connectivity analysis methods. In both analyses, the target ROIs were used as the seed regions to assess their effects on other ROIs under various conditions. The background ROIs were not used as the seed, since in half of the conditions, no stimulus was presented (no input visual signal) in background region. The PPI connectivity results are shown in Figure 7. Figure 7A shows the  $p$  values from permutation tests of PPI estimates for the interaction between the Alignment and Background. One significant interaction was found in tgV1 ROI when tgV2 was used as the seed (with permutation test,  $p = 0.042$ ). In detail, as shown in Figure 7B when the background was present, in tgV1 ROI, beta estimates for the aligned condition were larger than the unaligned ( $albg - uabg = 0.14$ , with permutation test  $p = 0.023$ ), but they were not different when the background was absent. Using a two-way ANOVA model (assumptions were satisfied based on the Lilliefors and

Bartlett's test) within each ROI (Alignment and Background as fixed effects, and subjects as a random effect), a weak interaction between Alignment and Background was observed in tgV1 ROI,  $F(1, 59) = 3.64$ ,  $p = 0.077$ . The correlation differences were also retinotopically specific to the target ROIs: no effect was observed in the background ROIs (see Figure A1 for the summary results from other ROIs when tgV2 was the seed).

With a second interarea connectivity analysis, the beta series correlations, a beta value was first estimated for each experimental trial. Next, these beta values were sorted according to conditions and correlated among ROIs for each condition. The acquired correlation coefficients for each condition and each seed and test ROI pair were tested for interaction between Alignment and Background factors (the permutation test  $p$  values as shown in Figure 8A). We found that, similar to the PPI results, the interaction was significant in tgV1 when tgV2 was the seed, with permutation test  $p = 0.021$  and with ANOVA test,  $F(1, 59) = 4.35$ ,  $p = 0.056$ . Figure 8B shows differences of beta series correlations estimated in tgV1 against tgV2 ROI (see Figure A2 for results in other ROIs). With the background, the correlation coefficients of beta values from the tgV1 when tgV2 was the seed were larger when the Gabors were aligned ( $albg - uabg = 0.079$ , with two-sided permutation test  $p = 0.0068$ ), but no significant difference was observed without background clutter. Due to symmetry of the analysis, the interaction in tgV2 when tgV1 was the seed was also significant. The beta series correlations showed similar sensitivity to tgV1–tgV2 connectivity as using the PPI analysis.

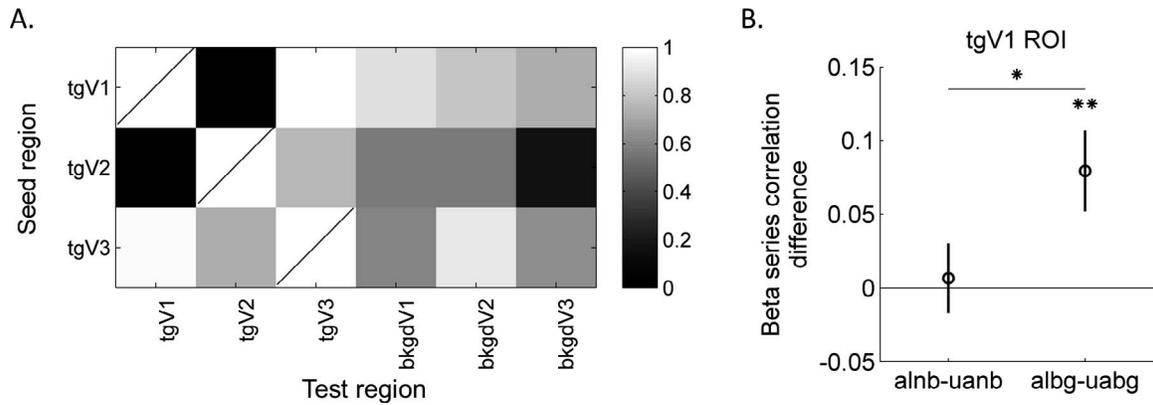


Figure 8. Connectivity results using beta series correlations. (A) The permutation test  $p$  values for interactions of beta series correlation coefficients between the Alignment and Background. A stronger interaction (small  $p$  value) is shown in dark color. A significant interaction was observed among tgV1 and tgV2 ROIs. (B) The differences of correlation coefficients among conditions in tgV1 are shown when tgV2 was the seed. The mean difference of beta series correlations from the 15 observers between the aligned and unaligned contours when there was no background is shown on the left, and the difference when there was background clutter was on the right. With background, the aligned condition tended to show larger correlation coefficients between the tgV1 and tgV2 ROIs. Error bars show  $\pm 1$  SE. Asterisks show significant levels based on the permutation test at  $*p < 0.05$  and  $**p < 0.01$ .

## Discussion

### Response enhancement in early visual areas is context dependent

We investigated how context and contour coherence affect the magnitude and interarea correlations of the fMRI BOLD signal in human early visual areas. The results from our first analysis showed that with clutter in the background, fMRI responses in target ROIs in early visual areas V1 and V2 were larger for aligned than for unaligned contours (Altmann et al., 2003; Kourtzi et al., 2003); while with isolated structure, the responses in V1 target ROIs were slightly larger for the unaligned than for aligned contours (Murray et al., 2002). This interaction between the contour alignment and the background context was significant in the regions retinotopically corresponding to the target stimulation in areas V1 and V2. Area V2 showed the most significant effect, which agrees with its role in extracting features from complex visual scenes (Boynton & Hegd , 2004; Huang, Hess, & Dakin, 2006; Ito & Komatsu, 2004; Merigan, Nealey, & Maunsell, 1993; Roe, 2003; von der Heydt, Peterhans, & Baumgartner, 1984). In contrast to previous functional imaging studies, we used a separate functional localizer to define the cortical region retinotopically associated with the target contour; we demonstrated that the effect was specific to the target related region, since no significant interaction was observed in the background ROIs.

Our results are consistent with the electrophysiological results reported in Li et al. (2006). They showed a close correlation between the responses of monkey V1

neurons and the perceptual saliency of contours, which was modulated by number of collinear elements or relative spacing between them; however, the correlation could be either positive or negative, depending on the context beyond the collinear elements. Specifically, they found that without the background clutter, neurons in V1 showed facilitation to three collinear lines compared with a single line in their receptive fields, but responses in V1 neurons were inhibited when there were more than three aligned elements. Conversely, with the background clutter, neuronal responses increased monotonically with increasing aligned line segments.

Similarly, we found that the increased response to aligned contours in the target ROIs relied on the presence of background clutter. This facilitative results to aligned contours observed in the presence of the background could be predicted based on flank facilitation on local segments along the contour. Both psychophysics and electrophysiology have shown that when surround segments are positioned within a certain range outside the neuron's receptive field and placed collinearly with the central stimulus, the neuronal responses to the center can be facilitated (Chen & Tyler, 2008; Kapadia et al., 1995; Kapadia, Westheimer, & Gilbert, 2000; Polat & Sagi, 1993, 1994). One interpretation is that the intrinsic horizontal connections in V1 can link neurons with nonoverlapping receptive fields but with similar orientation preference to integrate information over a relatively large visual field (Angelucci et al., 2002; Bosking, Zhang, Schofield, & Fitzpatrick, 1997; Gilbert, 1992; Gilbert & Wiesel, 1989; Li, 1998; Malach, Amir, Harel, & Grinvald, 1993; McGuire, Gilbert, Rivlin, & Wiesel, 1991; Rockland & Lund, 1983; Stettler, Das, Bennett, & Gilbert, 2002; Ts'o, Gilbert, & Wiesel, 1986). Therefore, the facili-

tated single neuron responses can be associated to form a coherent contour (Field et al., 1993; Hess & Field, 1999; Li & Gilbert, 2002), and cause response increases along the path. Furthermore, feedback signals from higher cortical areas, such as area V4, can also enhance the global contour signals in early visual areas (Chen et al., 2014).

In contrast, the nonfacilitative results may rely more on an understanding of the global scene. The visual system efficiently represents structures that follow natural scene statistics (Attneave, 1954; Barlow, 1961; Geisler, Perry, Super, & Gallogly, 2001; Simoncelli & Olshausen, 2001). Below, we discuss two ways the visual system might achieve this efficiency: predictive coding and disambiguation.

One way to achieve efficiency is to generate and feed back high-level “summary” templates for probable forms of the natural inputs, such as circular forms (Sigman, Cecchi, Gilbert, & Magnasco, 2001), and only signals representing deviations from the predicted templates are subsequently carried forward to be resolved by further processing, a theory referred to as “predictive coding” (Friston, 2005; MacKay, 1956; Mumford, 1992; Murray et al., 2004; Rao & Ballard, 1999). The degree of deviation from the template would be reflected in the magnitude of the *localized* neural activity (but see de-Wit, Kubilius, Wagemans, & Op de Beeck [2012]). In our experiment, predictive coding theory would result in larger responses to the unaligned contour because the elements deviate from the circular template.

Alternatively, the predictions from a higher level could disambiguate the lower-level representation by attenuating responses to unmatched incoming features (Murray et al., 2004; Yuille & Kersten, 2006). For example, the signal for background clutter could be suppressed once the circular foreground structure is detected, and this may decrease overall spatially averaged responses to coherent structure, perhaps for the purpose of metabolic efficiency (Barlow, 1959, 1961). Based on this model, when a coherent target appears, cortical responses to the target would be enhanced whereas responses to surrounding noise would be suppressed, as shown by Gilad et al. (2013) using voltage-sensitive dye imaging in V1 of monkeys.

However, based on our results, the response dependence on the background clutter was only observed at the regions specifically responding to the target ring and no suppression was observed in background ROIs, which are not consistent with the disambiguation theory described above. Although not statistically significant, the enhanced responses in V1 to unaligned contours (in the absence of background clutter) are consistent with the predictive coding idea.

## Coordination between areas V1 and V2

Furthermore, using the connectivity analyses—the psychophysiological interactions and beta series correlations—we found that the coordination between target V1 and target V2 ROIs was also highly dependent on the stimulus conditions. When the contours were presented together with the background clutter, a larger connectivity was observed between tgV1 and tgV2 ROIs for aligned contours than for unaligned ones. However, no connectivity difference was observed when the contours were presented without the background.

Close connections between areas V1 and V2 are well established from anatomical and physiological studies and can be grouped into three categories: feedforward, feedback, and common inputs. In the feedforward category, a large percentage of the cortical inputs in visual area V2 is from area V1 (Felleman & Van Essen, 1991; Sincich, Adams, & Horton, 2003), and the connections form multiple parallel pathways each carrying specific local representations such as color, form, and motion from V1 neurons (Federer et al., 2009; Livingstone & Hubel, 1988; Sincich & Horton, 2002, 2005). Area V1 also receives numerous feedback projections from V2 (Anderson & Martin, 2009; Angelucci et al., 2002; Barone, Batardiere, Knoblauch, & Kennedy, 2000; Girard, Hupé, & Bullier, 2001; Rockland & Virga, 1989; Stettler et al., 2002), but their function could be highly dependent on the visual stimulus (Anderson & Martin, 2009). For example, with one isolated stimulus no change in orientation selectivity was observed in area V1 when area V2 was cooled (Sandell & Schiller, 1982), and only a few V1 neurons were affected by V2 inactivation when simple center-surround stimuli were used (Bullier, Hupé, James, & Girard, 1996; Hupé, James, Girard, & Bullier, 2001). However, feedback may play a stronger role in more complex scenes. For example, given illusory contours induced by abutting gratings, there is evidence that area V2 modulates the orientation representation map in area V1 to provide a signature for a “higher order” contour (Ramsden, Hung, & Roe, 2001; Roe, 2003). The third category of connections that would relate responses in area V1 to V2 is the common inputs to these two areas from the same cortical and subcortical structures (Kennedy & Bullier, 1985). One example is that responses in areas V1 and V2 both are strongly influenced by feedback from area MT (Dakin, 2009; Hupé et al., 1998; Sillito, Cudeiro, & Jones, 2006).

Although the precise mechanism is unclear, the modulation of interarea correlations (aligned vs. unaligned) that we observed between V1 and V2, which were unique to the cluttered conditions, may reflect the concurrent signal transferring required to isolate the

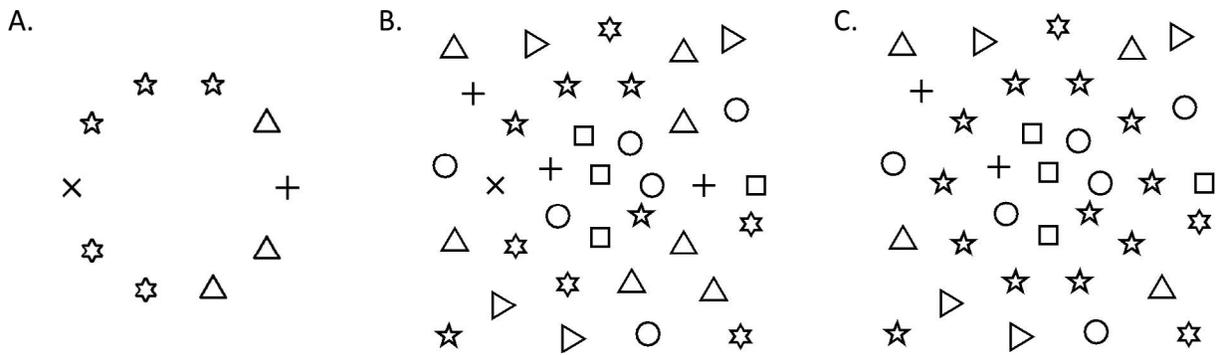


Figure 9. (A) A circular shape could be grouped based on global features. (B) The same circle is not easily identified when surrounded by clutter. Local linkage cues are required. For example, in (C), the nearby pentagrams serve as local cues for the circular shape to be grouped.

target contour from background noise. Similar inter-area correlation increases with increasing levels of elements alignment or structure detectability have been observed in previous studies (Cardin et al., 2010; Freeman, Donner, & Heeger, 2011).

### Flexible deployment of local versus global strategies balances accuracy and efficiency

Form information can be grouped based on either global or local features. Strategies based on global features (e.g., “elements form a circle”) may rely more on top-down templates: as shown in Figure 9A, a coherent circular shape could be perceived even when individual elements do not all share similar features. In contrast, a local feature-based strategy strongly depends on bottom-up relays, and it relies on similarities among nearby elements (Figure 9C); among clutter, when the nearby elements forming the circle do not share features, local integration processes could fail (Figure 9B; see also Keeble & Hess [1999], Levi & Klein [2000]). In fact, the prediction of human performance for detecting naturalistic contours among background distractors is fairly accurate by applying local grouping functions (Geisler et al., 2001). However, even when the local strategy is dominant, in order to compensate for the potential ambiguities of a merely bottom-up process, a higher-level global template can also be useful (Elder, Krupnik, & Johnston, 2003; Epshtein, Lifshitz, & Ullman, 2008). In all, applying both local and global strategies can be crucial to accurately represent the contour information (Friston, 2005; Kersten, Mamassian, & Yuille, 2004; Kersten & Yuille, 2014).

Further, the use of a global versus local strategy should be dynamically adjusted given the signal-to-noise ratio in the scene in order to improve efficiency in representations (Zhaoping, 2014). For example, in the absence of background clutter, once the system figures out the representation of a coherent circular contour,

no signal enhancement of the local alignment is necessary, which may explain the lack of facilitative responses to the coherent target in our study. Overall, balancing the above two strategies according to a greater stimulus context could improve both accuracy and efficiency of the cortical function.

## Conclusion

In summary, we have shown that the cortical responses of human early visual areas to coherent contours are affected by a larger context around them. A locally coherent target enhances neuronal responses to indicate certainty, but this may not be efficient especially when the target could be easily abstracted or explained globally. On the other hand, a system that only relies on the global approach and entirely discards basic representation signals from lower level areas is inflexible and may encounter problems later on, for example, when further operations on other detailed features are required. Therefore, the visual system should be able to apply both local and global strategies and to weight them according to context, such as scene complexity or task difficulty. With the current task, detecting circular contours among simple or complex scenes, we have found that early visual areas V1 and V2 may play an important role in manipulating contour integration strategies under various conditions.

*Keywords:* contour integration, fMRI, early visual areas

## Acknowledgments

We would like to thank D. Do, J. F. Schumacher, A. Grant, and M.-P. Schallmo for scanning assistance. We would also like to thank the reviewers for their

comments that helped improve the manuscript. This work was supported by the following funding/grant sources: NIH R21NS075525, P30NS076408, P41 EB015894, and ONR N000141210883.

Commercial relationships: none.

Corresponding author: Cheng Qiu.

Email: qiuxx077@umn.edu.

Address: Department of Psychology, University of Minnesota, Minneapolis, MN, USA.

## References

- Allman, J., Miezin, F., & McGuinness, E. (1985). Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local–global comparisons in visual neurons. *Annual Review of Neuroscience*, *8*, 407–430, doi:10.1146/annurev.ne.08.030185.002203.
- Altmann, C. F., Bühlhoff, H. H., & Kourtzi, Z. (2003). Perceptual organization of local elements into global shapes in the human visual cortex. *Current Biology*, *13*, 342–349, doi:10.1016/S0960-9822(03)00052-6.
- Anderson, J. C., & Martin, K. A. C. (2009). The synaptic connections between cortical areas V1 and V2 in macaque monkey. *The Journal of Neuroscience*, *29*, 11283–11293, doi:10.1523/JNEUROSCI.5757-08.2009.
- Angelucci, A., Levitt, J. B., Walton, E. J. S., Hupé, J.-M., Bullier, J., & Lund, J. S. (2002). Circuits for local and global signal integration in primary visual cortex. *The Journal of Neuroscience*, *22*, 8633–8646.
- Attneave, F. (1954). Some informational aspects of visual perception. *Psychological Review*, *61*, 183–193.
- Bandettini, P. A., Jesmanowicz, A., Wong, E. C., & Hyde, J. S. (1993). Processing strategies for time-course data sets in functional MRI of the human brain. *Magnetic Resonance in Medicine*, *30*, 161–173, doi:10.1002/mrm.1910300204.
- Barlow, H. B. (1959). Sensory mechanisms, the reduction of redundancy, and intelligence. *NPL Symposium on the Mechanization of Thought Process*, *10*, 535–539.
- Barlow, H. B. (1961). Possible principles underlying the transformation of sensory messages. In W. A. Rosenblith (Ed.), *Sensory Communication* (pp. 216–234). Cambridge, MA: The MIT Press.
- Barone, P., Batardiere, A., Knoblauch, K., & Kennedy, H. (2000). Laminar distribution of neurons in extrastriate areas projecting to visual areas V1 and V4 correlates with the hierarchical rank and indicates the operation of a distance rule. *The Journal of Neuroscience*, *20*, 3263–3281.
- Bauer, R., & Heinze, S. (2002). Contour integration in striate cortex. Classic cell responses or cooperative selection? *Experimental Brain Research*, *147*, 145–152, doi:10.1007/s00221-002-1178-6.
- Bosking, W. H., Zhang, Y., Schofield, B., & Fitzpatrick, D. (1997). Orientation selectivity and the arrangement of horizontal connections in tree shrew striate cortex. *The Journal of Neuroscience*, *17*, 2112–2127.
- Boynton, G. M., & Hegdé, J. (2004). Visual cortex: The continuing puzzle of area V2. *Current Biology*, *14*, R523–R524, doi:10.1016/j.cub.2004.06.044.
- Brainard, D. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.
- Bullier, J., Hupé, J. M., James, A., & Girard, P. (1996). Functional interactions between areas V1 and V2 in the monkey. *Journal of Physiology–Paris*, *90*, 217–220.
- Cardin, V., Friston, K. J., & Zeki, S. (2010). Top-down modulations in the visual form pathway revealed with dynamic causal modeling. *Cerebral Cortex*, *21*, 550–562, doi:10.1093/cercor/bhq122.
- Chen, C.-C., & Tyler, C. W. (2008). Excitatory and inhibitory interaction fields of flankers revealed by contrast-masking functions. *Journal of Vision*, *8*(4): 10, 1–14, doi:10.1167/8.4.10. [PubMed] [Article]
- Chen, M., Yan, Y., Gong, X., Gilbert, C. D., Liang, H., & Li, W. (2014). Incremental integration of global contours through interplay between visual cortical areas. *Neuron*, *82*, 682–694, doi:10.1016/j.neuron.2014.03.023.
- Cisler, J. M., Bush, K., & Steele, J. S. (2014). A comparison of statistical methods for detecting context-modulated functional connectivity in fMRI. *NeuroImage*, *84*, 1042–1052, doi:10.1016/j.neuroimage.2013.09.018.
- Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research*, *29*(3), 162–173, doi: http://dx.doi.org/10.1006/cbmr.1996.0014.
- Dakin, S. C. (2009). Vision: Thinking globally, acting locally. *Current Biology*, *19*, R851–R854, doi:10.1016/j.cub.2009.08.021.
- Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical surface-based analysis: I. Segmentation and surface reconstruction. *NeuroImage*, *9*, 179–194, doi:10.1006/nimg.1998.0395.
- de-Wit, L. H., Kubilius, J., Wagemans, J., & Op de

- Beeck, H. P. (2012). Bistable Gestalts reduce activity in the whole of V1, not just the retinotopically predicted parts. *Journal of Vision*, *12*(11): 12, 1–14, doi:10.1167/12.11.12. [PubMed] [Article]
- DeYoe, E. A., Carman, G. J., Bandettini, P., Glickman, S., Wieser, J., Cox, R., & Neitz, J. (1996). Mapping striate and extrastriate visual areas in human cerebral cortex. *Proceedings of the National Academy of Sciences, USA*, *93*, 2382–2386.
- Dumoulin, S. O., Dakin, S. C., & Hess, R. F. (2008). Sparsely distributed contours dominate extra-striate responses to complex scenes. *NeuroImage*, *42*, 890–901, doi:10.1016/j.neuroimage.2008.04.266.
- Dumoulin, S. O., & Hess, R. F. (2006). Modulation of V1 activity by shape: Image-statistics or shape-based perception? *Journal of Neurophysiology*, *95*, 3654–3664, doi:10.1152/jn.01156.2005.
- Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Boca Raton, FL: Chapman and Hall/CRC.
- Elder, J. H., Krupnik, A., & Johnston, L. A. (2003). Contour grouping with prior models. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*, 661–674, doi:10.1109/TPAMI.2003.1201818.
- Engel, S. A., Glover, G. H., & Wandell, B. A. (1997). Retinotopic organization in human visual cortex and the spatial precision of functional MRI. *Cerebral Cortex*, *7*, 181–192, doi:10.1093/cercor/7.2.181.
- Epshtein, B., Lifshitz, I., & Ullman, S. (2008). Image interpretation by a single bottom-up top-down cycle. *Proceedings of the National Academy of Sciences, USA*, *105*, 14298–14303, doi:10.1073/pnas.0800968105.
- Fang, F., Kersten, D., & Murray, S. O. (2008). Perceptual grouping and inverse fMRI activity patterns in human visual cortex. *Journal of Vision*, *8*(7):2, 1–9, doi:10.1167/8.7.2. [PubMed] [Article]
- Federer, F., Ichida, J. M., Jeffs, J., Schiessl, I., McLoughlin, N., & Angelucci, A. (2009). Four projection streams from primate V1 to the cytochrome oxidase stripes of V2. *The Journal of Neuroscience*, *29*, 15455–15471, doi:10.1523/JNEUROSCI.1648-09.2009.
- Felleman, D. J., & Van Essen, D. C. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral Cortex*, *1*(1), 1–47.
- Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: Evidence for a local “association field.” *Vision Research*, *33*, 173–193, doi:10.1016/0042-6989(93)90156-Q.
- Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis: II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, *9*(2), 195–207, doi:10.1006/nimg.1998.0396.
- Fitzpatrick, D. (2000). Seeing beyond the receptive field in primary visual cortex. *Current Opinion in Neurobiology*, *10*, 438–443.
- Freeman, J., Donner, T. H., & Heeger, D. J. (2011). Inter-area correlations in the ventral visual pathway reflect feature integration. *Journal of Vision*, *11*(4):15, 1–23, doi:10.1167/11.4.15. [PubMed] [Article]
- Friston, K. J. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *360*, 815–836, doi:10.1098/rstb.2005.1622.
- Friston, K. J., Buechel, C., Fink, G. R., Morris, J., Rolls, E., & Dolan, R. J. (1997). Psychophysiological and modulatory interactions in neuroimaging. *NeuroImage*, *6*, 218–229, doi:10.1006/nimg.1997.0291.
- Geisler, W. S., Perry, J. S., Super, B. J., & Gallogly, D. P. (2001). Edge co-occurrence in natural images predicts contour grouping performance. *Vision Research*, *41*, 711–724, doi:10.1016/S0042-6989(00)00277-7.
- Gilad, A., Meirovithz, E., & Slovin, H. (2013). Population responses to contour integration: Early encoding of discrete elements and late perceptual grouping. *Neuron*, *78*, 389–402, doi:10.1016/j.neuron.2013.02.013.
- Gilbert, C. D. (1992). Horizontal integration and cortical dynamics. *Neuron*, *9*(1), 1–13, doi:10.1016/0896-6273(92)90215-Y.
- Gilbert, C. D., & Wiesel, T. N. (1989). Columnar specificity of intrinsic horizontal and corticocortical connections in cat visual cortex. *The Journal of Neuroscience*, *9*, 2432–2442.
- Girard, P., Hupé, J. M., & Bullier, J. (2001). Feedforward and feedback connections between areas V1 and V2 of the monkey have similar rapid conduction velocities. *Journal of Neurophysiology*, *85*, 1328–1331.
- Gitelman, D. R., Penny, W. D., Ashburner, J., & Friston, K. J. (2003). Modeling regional and psychophysiological interactions in fMRI: The importance of hemodynamic deconvolution. *NeuroImage*, *19*, 200–207, doi:10.1016/S1053-8119(03)00058-2.
- Grill-Spector, K., Kourtzi, Z., & Kanwisher, N. (2001). The lateral occipital complex and its role in object recognition. *Vision Research*, *41*, 1409–1422, doi:10.1016/S0042-6989(01)00073-6.

- Grill-Spector, K., Kushnir, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). Cue-invariant activation in object-related areas of the human occipital lobe. *Neuron*, *21*, 191–202, doi:10.1016/S0896-6273(00)80526-7.
- Grill-Spector, K., Kushnir, T., Hendler, T., Edelman, S., Itzchak, Y., & Malach, R. (1998). A sequence of object-processing stages revealed by fMRI in the human occipital lobe. *Human Brain Mapping*, *6*, 316–328.
- Gross, C. G., Rocha-Miranda, C. E., & Bender, D. B. (1972). Visual properties of neurons in inferotemporal cortex of the macaque. *Journal of Neurophysiology*, *35*, 96–111.
- Händel, B., Lutzenberger, W., Thier, P., & Haarmer, T. (2007). Opposite dependencies on visual motion coherence in human area MT+ and early visual cortex. *Cerebral Cortex*, *17*, 1542–1549, doi:10.1093/cercor/bhl063.
- Harrison, L. M., Stephan, K. E., Rees, G., & Friston, K. J. (2007). Extra-classical receptive field effects measured in striate cortex with fMRI. *NeuroImage*, *34*, 1199–1208, doi:10.1016/j.neuroimage.2006.10.017.
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, *293*, 2425–2430, doi:10.1126/science.1063736.
- Hess, R. F., & Field, D. (1999). Integration of contours: New insights. *Trends in Cognitive Sciences*, *3*, 480–486, doi:10.1016/S1364-6613(99)01410-2.
- Huang, P.-C., Hess, R. F., & Dakin, S. C. (2006). Flank facilitation and contour integration: Different sites. *Vision Research*, *46*, 3699–3706, doi:10.1016/j.visres.2006.04.025.
- Hupé, J.-M., James, A. C., Girard, P., & Bullier, J. (2001). Response modulations by static texture surround in area V1 of the macaque monkey do not depend on feedback connections from V2. *Journal of Neurophysiology*, *85*, 146–163.
- Hupé, J.-M., James, A. C., Payne, B. R., Lomber, S. G., Girard, P., & Bullier, J. (1998). Cortical feedback improves discrimination between figure and background by V1, V2, and V3 neurons. *Nature*, *394*, 784–787, doi:10.1038/29537.
- Ito, M., & Komatsu, H. (2004). Representation of angles embedded within contour stimuli in area V2 of macaque monkeys. *Journal of Neuroscience*, *24*, 3313–3324, doi:10.1523/JNEUROSCI.4364-03.2004.
- Kapadia, M. K., Ito, M., Gilbert, C. D., & Westheimer, G. (1995). Improvement in visual sensitivity by changes in local context: Parallel studies in human observers and in V1 of alert monkeys. *Neuron*, *15*, 843–856, doi:10.1016/0896-6273(95)90175-2.
- Kapadia, M. K., Westheimer, G., & Gilbert, C. D. (2000). Spatial distribution of contextual interactions in primary visual cortex and in visual perception. *Journal of Neurophysiology*, *84*, 2048–2062.
- Keeble, D. R. T., & Hess, R. F. (1999). Discriminating local continuity in curved figures. *Vision Research*, *39*, 3287–3299, doi:10.1016/S0042-6989(99)00021-8.
- Kennedy, H., & Bullier, J. (1985). A double-labeling investigation of the afferent connectivity to cortical areas V1 and V2 of the macaque monkey. *The Journal of Neuroscience*, *5*, 2815–2830.
- Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology*, *55*, 271–304, doi:10.1146/annurev.psych.55.090902.142005.
- Kersten, D., & Yuille, A. (2014). Vision: Bayesian inference and beyond. In J. S. Werner & L. M. Chalupa (Eds.), *The new visual neurosciences* (pp. 1263–1278). Cambridge, MA: The MIT Press.
- Kim, J., & Horwitz, B. (2008). Investigating the neural basis for fMRI-based functional connectivity in a blocked design: Application to interregional correlations and psycho-physiological interactions. *Magnetic Resonance Imaging*, *26*, 583–593, doi:10.1016/j.mri.2007.10.011.
- Kourtzi, Z., & Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *The Journal of Neuroscience*, *20*, 3310–3318.
- Kourtzi, Z., Tolias, A. S., Altmann, C. F., Augath, M., & Logothetis, N. K. (2003). Integration of local features into global shapes: Monkey and human fMRI studies. *Neuron*, *37*, 333–346, doi:10.1016/S0896-6273(02)01174-1.
- Kovács, I. (1996). Gestalten of today: Early processing of visual contours and surfaces. *Behavioural Brain Research*, *82*(1), 1–11, doi:10.1016/S0166-4328(97)81103-5.
- Kovács, I., & Julesz, B. (1993). A closed curve is much more than an incomplete one: Effect of closure in figure-ground segmentation. *Proceedings of the National Academy of Sciences, USA*, *90*, 7495–7497.
- Kovács, I., & Julesz, B. (1994). Perceptual sensitivity maps within globally defined visual shapes. *Nature*, *370*, 644–646, doi:10.1038/370644a0.
- Lerner, Y., Hendler, T., Ben-Bashat, D., Harel, M., & Malach, R. (2001). A hierarchical axis of object processing stages in the human visual cortex.

- Cerebral Cortex*, 11, 287–297, doi:10.1093/cercor/11.4.287.
- Levi, D. M., & Klein, S. A. (2000). Seeing circles: What limits shape perception? *Vision Research*, 40, 2329–2339, doi:10.1016/S0042-6989(00)00092-4.
- Li, W., & Gilbert, C. D. (2002). Global contour saliency and local colinear interactions. *Journal of Neurophysiology*, 88, 2846–2856, doi:10.1152/jn.00289.2002.
- Li, W., Piëch, V., & Gilbert, C. D. (2006). Contour saliency in primary visual cortex. *Neuron*, 50, 951–962, doi:10.1016/j.neuron.2006.04.035.
- Li, Z. (1998). A neural model of contour integration in the primary visual cortex. *Neural Computation*, 10, 903–940, doi:10.1162/089976698300017557.
- Lilliefors, H. W. (1967). On the Kolmogorov-Smirnov test for normality with mean and variance unknown. *Journal of the American Statistical Association*, 62(318), 399–402, doi:10.1080/01621459.1967.10482916.
- Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: Anatomy, physiology, and perception. *Science*, 240, 740–749.
- Logothetis, N. K., & Sheinberg, D. L. (1996). Visual object recognition. *Annual Review of Neuroscience*, 19, 577–621, doi:10.1146/annurev.ne.19.030196.003045.
- MacKay, D. M. (1956). The epistemological problem for automata. In C. E. Shannon & J. McCarthy (Eds.), *Automata studies* (pp. 235–251). Princeton, NJ: Princeton University Press.
- Malach, R., Amir, Y., Harel, M., & Grinvald, A. (1993). Relationship between intrinsic connections and functional architecture revealed by optical imaging and in vivo targeted biocytin injections in primate striate cortex. *Proceedings of the National Academy of Sciences, USA*, 90, 10469–10473.
- McGuire, B. A., Gilbert, C. D., Rivlin, P. K., & Wiesel, T. N. (1991). Targets of horizontal connections in macaque primary visual cortex. *The Journal of Comparative Neurology*, 305, 370–392, doi:10.1002/cne.903050303.
- McKeefry, D. J., Watson, J. D. G., Frackowiak, R. S. J., Fong, K., & Zeki, S. (1997). The activity in human areas V1/V2, V3, and V5 during the perception of coherent and incoherent motion. *NeuroImage*, 5(1), 1–12, doi:10.1006/nimg.1996.0246.
- McLaren, D. G., Ries, M. L., Xu, G., & Johnson, S. C. (2012). A generalized form of context-dependent psychophysiological interactions (gPPI): A comparison to standard approaches. *NeuroImage*, 61, 1277–1286, doi:10.1016/j.neuroimage.2012.03.068.
- McManus, J. N. J., Li, W., & Gilbert, C. D. (2011). Adaptive shape processing in primary visual cortex. *Proceedings of the National Academy of Sciences, USA*, 108(24), 9739–9746, doi:10.1073/pnas.1105855108.
- Mendola, J. D., Dale, A. M., Fischl, B., Liu, A. K., & Tootell, R. B. H. (1999). 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.
- Merigan, W. H., Nealey, T. A., & Maunsell, J. H. (1993). Visual effects of lesions of cortical area V2 in macaques. *The Journal of Neuroscience*, 13, 3180–3191.
- Mumford, D. (1992). On the computational architecture of the neocortex: The role of cortico-cortical loops. *Biological Cybernetics*, 66, 241–251.
- Murray, S. O., Kersten, D., Olshausen, B. A., Schrater, P., & Woods, D. L. (2002). Shape perception reduces activity in human primary visual cortex. *Proceedings of the National Academy of Sciences, USA*, 99, 15164–15169, doi:10.1073/pnas.192579399.
- Murray, S. O., Schrater, P., & Kersten, D. (2004). Perceptual grouping and the interactions between visual cortical areas. *Neural Networks*, 17, 695–705, doi:10.1016/j.neunet.2004.03.010.
- Paradis, A. L., Cornilleau-Pérès, V., Droulez, J., Van De Moortele, P. F., Lobel, E., Berthoz, A., & Poline, J. B. (2000). Visual perception of motion and 3-D structure from motion: An fMRI study. *Cerebral Cortex*, 10, 772–783, doi:10.1093/cercor/10.8.772.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision*, 10, 437–442.
- Pettet, M. W. (1999). Shape and contour detection. *Vision Research*, 39, 551–557, doi:10.1016/S0042-6989(98)00130-8.
- Pettet, M. W., McKee, S. P., & Grzywacz, N. M. (1998). Constraints on long range interactions mediating contour detection. *Vision Research*, 38, 865–879, doi:10.1016/S0042-6989(97)00238-1.
- Polat, U., & Sagi, D. (1993). Lateral interactions between spatial channels: Suppression and facilitation revealed by lateral masking experiments. *Vision Research*, 33, 993–999, doi:10.1016/0042-6989(93)90081-7.
- Polat, U., & Sagi, D. (1994). The architecture of

- perceptual spatial interactions. *Vision Research*, *34*, 73–78, doi:10.1016/0042-6989(94)90258-5.
- Ramsden, B., Hung, C., & Roe, A. W. (2001). Real and illusory contour processing in area V1 of the primate: A cortical balancing act. *Cerebral Cortex*, *11*, 648–665, doi:10.1093/cercor/11.7.648.
- Rao, R. P. N., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*, 79–87.
- Rissman, J., Gazzaley, A., & D’Esposito, M. (2004). Measuring functional connectivity during distinct stages of a cognitive task. *NeuroImage*, *23*, 752–763, doi:10.1016/j.neuroimage.2004.06.035.
- Rockland, K. S., & Lund, J. S. (1983). Intrinsic laminar lattice connections in primate visual cortex. *The Journal of Comparative Neurology*, *216*, 303–318, doi:10.1002/cne.902160307.
- Rockland, K. S., & Virga, A. (1989). Terminal arbors of individual “Feedback” axons projecting from area V2 to V1 in the macaque monkey: A study using immunohistochemistry of anterogradely transported *Phaseolus vulgaris*-leucoagglutinin. *The Journal of Comparative Neurology*, *285*, 54–72, doi:10.1002/cne.902850106.
- Roe, A. W. (2003). Modular complexity of area V2 in the macaque monkey. In C. Collins & J. Kaas (Eds.), *The primate visual system* (pp. 109–138). New York: CRC Press.
- Roelfsema, P. R., Lamme, V. A. F., & Spekreijse, H. (2004). Synchrony and covariation of firing rates in the primary visual cortex during contour grouping. *Nature Neuroscience*, *7*, 982–991, doi:10.1038/nn1304.
- Sandell, J. H., & Schiller, P. H. (1982). Effect of cooling area 18 on striate cortex cells in the squirrel monkey. *Journal of Neurophysiology*, *48*, 38–48.
- Sereno, M. I., Dale, A. M., Reppas, J. B., Kwong, K. K., Belliveau, J. W., Brady, T. J., & Tootell, R. B. (1995). Borders of multiple visual areas in humans revealed by functional magnetic resonance imaging. *Science*, *268*, 889–893.
- Sigman, M., Cecchi, G. A., Gilbert, C. D., & Magnasco, M. O. (2001). On a common circle: Natural scenes and Gestalt rules. *Proceedings of the National Academy of Sciences, USA*, *98*, 1935–1940.
- Sillito, A. M., Cudeiro, J., & Jones, H. E. (2006). Always returning: Feedback and sensory processing in visual cortex and thalamus. *Trends in Neurosciences*, *29*, 307–316, doi:10.1016/j.tins.2006.05.001.
- Simoncelli, E. P., & Olshausen, B. A. (2001). Natural image statistics and neural representation. *Annual Review of Neuroscience*, *24*, 1193–1216.
- Sincich, L. C., Adams, D. L., & Horton, J. C. (2003). Complete flatmounting of the macaque cerebral cortex. *Visual Neuroscience*, *20*, 663–686.
- Sincich, L. C., & Horton, J. C. (2002). Divided by cytochrome oxidase: A map of the projections from V1 to V2 in macaques. *Science*, *295*, 1734–1737.
- Sincich, L. C., & Horton, J. C. (2005). The circuitry of V1 and V2: Integration of color, form, and motion. *Annual Review of Neuroscience*, *28*, 303–326, doi:10.1146/annurev.neuro.28.061604.135731.
- Smith, S. M., Jenkinson, M., Woolrich, M. W., Beckmann, C. F., Behrens, T. E. J., Johansen-Berg, H., & Matthews, P. M. (2004). Advances in functional and structural MR image analysis and implementation as FSL. *NeuroImage*, *23 Suppl. 1*, S208–S219, doi:10.1016/j.neuroimage.2004.07.051.
- Snedecor, G. W., & Cochran, W. G. (1956). *Statistical methods*. Ames, IA: Iowa State University Press.
- Stettler, D. D., Das, A., Bennett, J., & Gilbert, C. D. (2002). Lateral connectivity and contextual interactions in macaque primary visual cortex. *Neuron*, *36*, 739–750.
- Tanaka, K. (1996). Inferotemporal cortex and object vision. *Annual Review of Neuroscience*, *19*, 109–139, doi:10.1146/annurev.ne.19.030196.000545.
- Tanskanen, T., Saarinen, J., Parkkonen, L., & Hari, R. (2008). From local to global: Cortical dynamics of contour integration. *Journal of Vision*, *8(7):15*, 1–12, doi:10.1167/8.7.15. [PubMed] [Article]
- Ts’o, D. Y., Gilbert, C. D., & Wiesel, T. N. (1986). Relationships between horizontal interactions and functional architecture in cat striate cortex as revealed by cross-correlation analysis. *The Journal of Neuroscience*, *6*, 1160–1170.
- Ullman, S., & Sha’ashua, A. (1988, December). Structural saliency: The detection of globally salient structures using a locally connected network. Paper presented at *Computer Vision., Second International Conference on*, Tampa, FL. doi:10.1109/CCV.1988.590008.
- von der Heydt, R., Peterhans, E., & Baumgartner, G. (1984). Illusory contours and cortical neuron responses. *Science*, *224*, 1260–1262, doi:10.1126/science.6539501.
- Yuille, A., & Kersten, D. (2006). Vision as Bayesian inference: Analysis by synthesis? *Trends in Cognitive Sciences*, *10*, 301–308, doi:10.1016/j.tics.2006.05.002.
- Zhaoping, L. (2014). *Understanding vision: Theory, models, and data*. New York, NY: Oxford University Press.

## Appendix

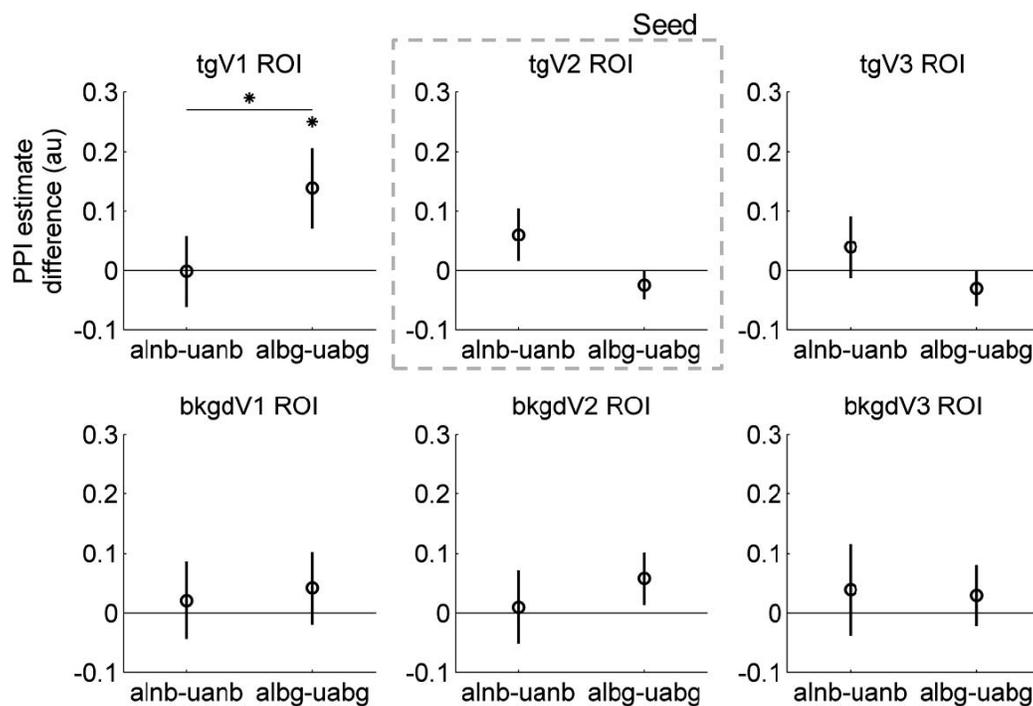


Figure A1. Connectivity results in test ROIs using the PPI analysis when tgV2 was the seed ROI (dashed frame). Each panel shows the average results from the 15 observers in one ROI. Error bars show  $\pm 1$  SE. In tgV1 ROI (the 1<sup>st</sup> panel), with background clutter (on the right), the aligned condition has larger PPI connectivity estimates than the unaligned condition. The results in other ROIs are not different from zero.

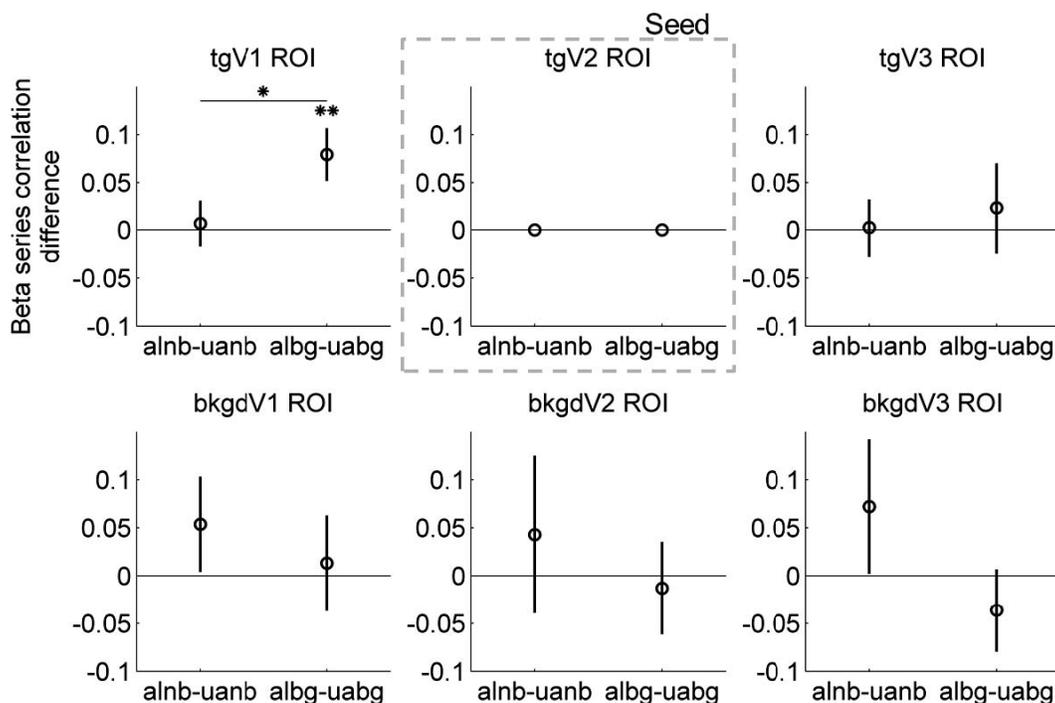


Figure A2. Connectivity results in test ROIs using the beta series method when tgV2 was the seed region (dashed frame). Each panel shows the average results from the 15 observers in one ROI. Error bars show  $\pm 1$  SE. In tgV1 ROI (the 1<sup>st</sup> panel), with background clutter (on the right), the aligned condition has larger beta series correlation coefficients than the unaligned condition. The results in other ROIs are not different from zero.