Contextual effects in human visual cortex depend on surface structure

Sung Jun Joo\textsuperscript{1,2,*} & Scott O. Murray\textsuperscript{1}

\textsuperscript{1}Department of Psychology, University of Washington, Seattle, WA, USA 98195
\textsuperscript{2}Center for Perceptual Systems and Department of Psychology, University of Texas, Austin, TX, USA 78712

Keywords: Contextual modulation, Contrast adaptation, Early visual cortex, ERP, Grouping,

Running head: Contextual effects depend on surface structure

*Correspondence should be addressed to Sung Jun Joo at the above address. E-mail: sjjoo@utexas.edu
Neural responses in early visual cortex depend on stimulus context. One of the most well-established context dependent effects is orientation-specific surround suppression: the neural response to a stimulus inside the receptive field of a neuron (“target”) is suppressed when it is surrounded by iso-oriented compared to orthogonal stimuli (“flankers”). Despite the importance of orientation-specific surround suppression in potentially mediating a number of important perceptual effects including saliency, contour integration, and orientation discrimination, the underlying neural mechanisms remain unknown. The suppressive signal could be inherited from pre-cortical areas as early as the retina and thalamus, arise from local circuits through horizontal connections, or be fed back from higher visual cortex. Here we show using two different methodologies—measurements of scalp-recorded event-related potentials (ERPs) and behavioral contrast adaptation aftereffects in humans—that orientation-specific surround suppression is dependent on the surface structure in an image. When the target and flankers can be grouped on the same surface (independent of their distance) orientation-specific surround suppression occurs. When the target and flankers are on different surfaces (independent of their distance) orientation-specific surround suppression does not occur. Our results demonstrate a surprising role of high-level, global processes such as grouping in determining when contextual effects occur in early visual cortex.

It is well established that there can be neural suppression in early visual cortex measured at the individual neuron level (Blakemore and Tobin, 1972; Maffei and Fiorentini, 1976; Allman et al., 1985; DeAngelis et al., 1994; Sillito et al., 1995; Cavanaugh et al., 2002), by human fMRI (Zenger-Landolt and Heeger, 2003; Joo et al., 2012), and by human ERP (Haynes et al., 2003; Joo et al., 2012) when surrounding stimuli (“flankers”) match the center (“target”) orientation: orientation-specific surround suppression. Despite the importance of orientation-specific surround suppression in potentially mediating a number of important perceptual effects including saliency (Knierim and Van Essen, 1992; Zipser et al., 1996; Kastner et al., 1997), contour integration (Dobbins et al., 1987; Kapadia et al., 1995), and orientation discrimination (Mareschal et al., 2001), the underlying neural mechanisms remain unknown. For example, the suppressive signal could be inherited from pre-cortical areas as early as the retina (Solomon et al., 2006) and thalamus (Alitto and Usrey, 2008), arise from V1 local circuits through horizontal connections (Kapadia et al., 2000; Adesnik et al., 2012), or be fed back from higher visual cortex (Angelucci et al., 2002; Bair et al., 2003).

However, a number of recent psychophysical findings have shown that high-level perceptual grouping can influence basic visual detection and discrimination performance in stimulus configurations that resemble those used in orientation-specific surround suppression experiments (Mareschal et al., 2001; Sayim et al., 2008; Huang et al., 2012; Joo et al., 2012; Manassi et al., 2012). Thus, we hypothesized that perceptual grouping would influence orientation-specific surround suppression such that it would occur only when the target and flankers were grouped into a single array. For example, we predicted that stimulus manipulations that isolated the target from the flankers—such as increasing the spatial separation or making the target and flankers appear to be on different surfaces—would eliminate orientation-specific surround suppression. On the other hand, manipulations that grouped the target and flankers into a single array—such as decreasing the spatial separation or
making the target and flankers appear to be on the same surface—would promote orientation-specific surround suppression.

Here we used two different experimental methodologies—behavioral contrast adaptation aftereffect and event-related potential (ERP) measurements—to test our grouping hypothesis. We measured the response to a target in stimulus configurations that manipulated distance (near and far) and surface placement (target and flankers on the same versus different surface). Our results demonstrate that orientation-specific surround suppression in early visual cortex occurs under stimulus conditions that promote grouping of the target and flankers into a single array of elements.

Materials and Methods

Observers

Experiment 1 (Behavioral contrast adaptation experiments): All observers had normal or corrected-to-normal vision and all gave informed written consent in accordance with the University of Texas at Austin Institutional Review Board. Five observers including the first author voluntarily participated.

Experiment 2 (ERP experiments): All observers had normal or corrected-to-normal vision and all gave informed written consent in accordance with the University of Washington Institutional Review Board. Eighteen observers including the first author participated in Experiment 2A. A total of 48 observers participated in Experiments 2B, 2C and 2D (16 per each Experiment). All observers were assigned to only one experiment. Except for the first author, all were naïve observers who volunteered for either course credit or monetary compensation ($20/hour).

Stimuli and procedure: Experiment 1

The stimuli were generated using MATALB PsychToolbox (Brainard, 1997; Pelli, 1997) on a Mac Pro PC and displayed on a 17-in SONY OLED monitor (PVM-1741, 60 Hz refresh rate with a resolution of 1080p and flicker-free mode on). The monitor was linearized using a standard gamma correction procedure. Ten-bit luminance steps were generated using the native 10-bit mode supported by the graphics card (ATI Radeon HD 4870) and PsychToolbox driver. The viewing distance was 68 cm.

The Gabor patches had a standard deviation of 0.72°, a spatial frequency of 2 cycles per degree (cpd), and were 30% contrast. The surface patches were generated using the “drop shadow” function (Mode: Overlay, Opacity: 75%, Blur: 0.18 cm) in Adobe Illustrator CS3 (Adobe Systems, Inc.). We manipulated X and Y offset parameters in Adobe Illustrator’s “drop shadow” function to minimize the spatial overlap between the target and the shadow in Experiments 1D and 2D. Figs. 1A and 1B show cross-sectional luminance profiles of these stimuli. A fixation point (0.48° in diameter) was displayed at the center of the display. The target is displaced by 6° horizontally from the fixation point.

Four stimulus configurations were used: (1) in Experiment 1A (near, no-surface configuration) the flankers were positioned so that the center-center distance of the target and flankers was 3°, (2) in Experiment 1B (far, no-surface configuration) the flankers were positioned so that the center-center distance was 6°, (3) in Experiment 1C (far, same-surface configuration) the target and flankers were separated by 6° and appeared on the same
rendered surface and (4) in Experiment 1D (near, different-surface configuration) the target and flankers were separated by 3° and rendered on different surfaces. In this condition, we adjusted the shadow position and size to prevent overlap between the shadow edge and the target.

To measure the contrast detection threshold for a target, three randomly interleaved independent QUEST (Watson and Pelli, 1983) staircases were used. The detection task was a two-interval forced choice (2IFC) task where observers indicated which interval had the target. Each interval (200 ms) was indicated by a high-pitched tone and there was a gray blank (300 ms) between intervals. Auditory feedback was given for an incorrect response. Each observer participated in 8 sessions (4 Experiments x 2 repetitions). We measured the amount of adaptation with a vertical target orientation embedded in either horizontal-orientation (“orthogonal” condition) or vertical-orientation (“same” condition) flankers. At the beginning of each session, each observer’s contrast detection threshold of the target was measured using the same 2IFC tasks to define the baseline. A session consisted of two adaptation blocks (“same” and “orthogonal” condition) and the order of blocks was counterbalanced across sessions per each observer. A five-minute break was inserted between blocks to prevent any carry-over effect from the previous block. There were 60 trials (20 trials per each staircase) in a block. The last contrast values of six staircases (2 repetitions x 3 staircases) were averaged to estimate an observer’s contrast detection threshold for 82% performance. If a staircase did not converge (greater than one standard deviation from the mean of the six data points) the staircase was discarded. This excluded 10 data points from a total possible of 240 data points across observers.

Adapting stimuli consisted of a target and flankers. Observers were initially adapted for 30 s, followed by the first 2-IFC task trial. A 5 s top-up adaptation period was inserted between subsequent trials to maintain stable adaptation (Fig. 1C). Stimuli were counter-phase flickered at 2 Hz. A 500 ms gray blank was inserted before each trial began. During this blank period, a black line was displayed next to the fixation point to indicate the beginning of a trial. The target location was always marked during both the adaptation and task periods to remove effects of location uncertainty on the detection task (Petrov et al., 2006). To equate the attentional state across conditions observers performed a contrast decrement task on the fixation mark during adaptation periods (Bi et al., 2009). The contrast decrement (10%) was displayed for 150 ms and the onset of the contrast decrement was selected randomly from a uniform distribution between 1 and 1.5 s. To quantify the amount of adaptation we defined the threshold ratio between detection threshold before and after adaptation (threshold$_{after}$/threshold$_{before}$).

**Stimuli and procedure: Experiment 2**

We used the same stimulus parameters and configurations as in Experiment 1 except for the following: (1) the fixation point was placed 3° below from the center of the display and (2) the target stimuli were displayed in the upper quadrant of both visual fields 3° horizontally and 3° vertically from the fixation point. The center-to-center distance between the stimuli comprising a pattern was 3° in Experiment 2A (the near, no-surface configuration) and Experiment 2D (the near, different-surface configuration) and 6° in Experiment 2B (the far, no-surface configuration) and Experiment 2C (the far, same-surface configuration). The stimuli were generated and controlled by Presentation (Neurobehavioral Systems, Inc.) on a PC, and
they were displayed on a 21-in. CRT monitor (60 Hz refresh rate). The viewing distance was approximately 70 cm.

Target orientation could be either vertical or horizontal orientation. There were three flanker conditions (“single”, “same”, and “orthogonal”). The flanker orientation varied according to the target orientation (see Fig. 2A). The flanker orientation matched the target orientation in the “same” condition but was orthogonal to the target orientation in the “orthogonal” condition. A thin circle (0.2”) that matched the size and contrast of Gabor stimuli was displayed in the flanker positions in the “single” condition to equate the stimulus timing.

Fig. 2B shows an example trial. On a given trial, flankers (or flankers and cast-shadow in Experiments 2C and 2D) appeared before the onset of the target. After a random duration chosen from a uniform distribution between 1 and 2 s, targets were briefly flashed for 100 ms. After the target offset, the flankers remained in the display for 500 ms. The inter-trial interval was 3 s. Observers were asked to maintain fixation and to limit eye blinks to the inter-trial interval.

One experimental block consisted of 12 trials (2 target orientation conditions x 3 flanker conditions x 2 repetitions). The order of trials was randomized within a block. Observers finished 22-51 blocks (264-612 trials). Observers initiated each block after a 5 s break by pressing a designated key on the button box. The first block served as practice.

EEG recording and data analysis

EEG waveforms were recorded using BioSemi active Ag-AgCl electrodes from 64 sites. The signals were referenced to the left mastoid during online acquisition and re-referenced to the average of right and left mastoids offline. Vertical EOG was measured using an electrode placed below the left eye and horizontal EOG was measured using an electrode placed at the outer canthus of the right eye. The signals were digitized at a sampling rate of 256 Hz.

EEG epochs started 100 ms before the target onset and lasted 400 ms after the target onset. Each waveform was baseline corrected to the average voltage of the interval -100 ms to 0 ms before the target onset and low-pass filtered at 40 Hz to remove high-frequency noise.

Trials with waveforms that had a larger than 50 µV peak-to-peak vertical and that exceeded ±50 µV on other electrodes were excluded as these were trials deemed to be contaminated with eye blinks or other sources of noise. Data from 5 observers (Experiment 2A) and 1 observer (Experiment 2B, 2C, and 2D) were discarded due to excessive artifact rejection (> 50%). The resulting waveforms were averaged across conditions individually for statistical analyses and then averaged across observers for figures.

P1 amplitude on six electrodes (Oz, O1, O2, POz, PO3, and PO4) was measured by averaging the ERP amplitudes during the time window of 130 ms to 170 ms. These electrodes were centered over the maximum of the P1 component (150 ms after target onset) as determined through visual inspection of the scalp topography (see Fig. 2C). These individual amplitudes were averaged across six electrodes to represent P1 amplitude. We conducted a repeated-measures ANOVA for the statistical analysis. The data from the “single” condition was not included in the analysis because the response to the “single” condition (interaction between an oriented target and circles) was categorically different from the other conditions (interaction between an oriented target and oriented flankers). The “single” condition was only
used to assess any difference between vertical and horizontal targets in the absence of oriented flankers and in the presence of the cast-shadows.

Results

Experiment 1: Behavioral Contrast Adaptation Experiments

We measured the amount of behavioral contrast adaptation to a vertical target Gabor with horizontal (orthogonal) flankers and vertical (same-orientation) flankers. Psychophysical contrast adaptation aftereffects can be used to infer the magnitude of the neural response in early visual cortex to the target stimulus (Blakemore and Campbell, 1969; Movshon and Lennie, 1979; Bradley et al., 1988). To quantify adaptation strength, we calculated the ratio of each observer’s contrast detection threshold for a target before and after adaptation. The assumption is that more adaptation—as indexed by an increase in post-adaptation detection thresholds—reflects stronger neural activity in response to the adapting stimulus (Blakemore and Campbell, 1969; Carandini et al., 1998; Dragoi et al., 2000; Kohn and Movshon, 2003; Engel, 2005; Fang et al., 2005; Blake et al., 2006; Larsson et al., 2006). To equate attentional state across conditions, subjects performed a demanding luminance decrement task at fixation. The mean performance on the fixation task was 90±7% and there were no significant performance differences across conditions.

We found adaptation aftereffects suggesting orientation-specific surround suppression only with stimulus configurations that promoted grouping of the target and flankers into a single array (Fig. 3A-D). First, in Experiment 1A we confirmed that the amount of adaptation was modulated by the orientation of nearby (3° separation) flankers. Specifically, the vertical target surrounded by nearby (3° separation) vertical flankers resulted in a smaller threshold ratio compared to nearby horizontal flankers (Fig. 3A; $t_4 = 3.78, p = 0.02$). This smaller threshold ratio in the “same” condition compared to the “orthogonal” condition is consistent with less neural activity to the target and a signature of orientation-specific surround suppression.

After establishing orientation-specific surround suppression using our adaptation protocol, we tested our grouping hypothesis. First, in Experiment 1B we simply increased the distance between the target and flankers by doubling the center-to-center distance (6°, Fig. 3B) used in Experiment 1A. Perceptually, with the increased distance, the stimulus now appeared to be 3 isolated Gabor patches rather than a single array of 3 Gabor patches following the well-known Gestalt principle of proximity. Thus, because the target and flankers were no longer grouped into a single array, we predicted that orientation-specific surround suppression would be eliminated. Consistent with this prediction, the basic orientation-specific surround suppression effect was not present with distant (6° separation) flankers (Fig. 3B; $t_4 = 0.06, p = 0.96$). However, this finding that increasing the distance between the target and flankers eliminates orientation-specific surround suppression is consistent with a number of potential explanations—ranging from local normalization models (Cavanaugh et al., 2002; Shapley, 2004) to our high-level, grouping hypothesis.

To distinguish between these alternatives, in Experiment 2C we used the distant-flanker configuration of Experiment 2B but made the target and flankers appear to be grouped on a common surface that was distinct from the background. This was done by adding a small “cast shadow” in the region around the stimuli to create a surface that appeared to be at a closer depth-plane than the background (Fig. 3C). Any model of orientation-specific surround...
suppression that emphasized local orientation interactions between the target and flankers would again predict no orientation-specific surround suppression—as in Experiment 1B. However, based on our grouping hypothesis, we expected to observe orientation-specific surround suppression because the target and flankers were now grouped on a common surface separate from the background. Consistent with the grouping hypothesis, when the distant-flanker configuration was displayed on a cast-shadow surface, orientation-specific surround suppression was restored: the threshold ratio was smaller in the “same” condition compared to the “orthogonal” condition (Fig. 3C; $t_4 = 3.40, p = 0.03$).

If surface representations are indeed important for determining when orientation-specific surround suppression effects occur, we predicted that moving the target and flankers to different surfaces would eliminate orientation-specific surround suppression—even with spatial parameters that would otherwise result in strong orientation-specific surround suppression. In Experiment 1D, we used the same spatial parameters of Experiment 1A—where we observed strong orientation-specific surround suppression—but moved the flankers to different surfaces than the target (Fig. 3D). Although the flankers were displayed in the near proximity of the target, orientation-specific surround suppression was eliminated: there was no difference in the amount of adaptation between the “same” and “orthogonal” condition (Fig. 3D; $t_4 = 0.40, p = 0.71$).

These findings suggest that low-level, long-term contrast adaptation can be modulated by high-level image structure while attention was controlled. In the next series of experiments, we used an ERP technique to measure the visual evoked potential (VEP) to the onset of the target to characterize the timecourse of the orientation-specific surround suppression in early visual cortex. To more generalize our findings, we used two target orientations (vertical and horizontal).

**Experiment 2: ERP Experiments**

We measured the ERP response to an oriented Gabor (“target”) with flankers above and below the target. Critically, to ensure that the flanker (and cast shadow) onset and offset did not contaminate the evoked potential to the target, flankers (and cast shadows) were displayed before target onset and remained in the display until after target offset (Fig. 2B and Materials and Methods). The target was briefly flashed for 100 ms. The duration between flanker onset and target onset was randomized between 1 and 2 s. This duration was long enough such that the visually evoked potentials to the flankers diminished to baseline levels before target onset thus ensuring that we measured the ERP response to only the target stimulus. The target could be either vertically or horizontally oriented and be displayed with flankers that matched the target orientation (“same” condition; Fig. 2A, left) or flankers that were orthogonally oriented compared to the target (“orthogonal” condition; Fig. 2A, right).

In Experiment 2A we established the basic orientation-specific surround suppression effect in the ERP response using stimulus configurations where the target was surrounded by “nearby” flankers (Fig. 4A, target-flanker distance = 3°). We used the amplitude of the earliest component of our data (P1; 150 ms after target onset) to index neural activity in early visual cortex (Clark et al., 2004; Joo et al., 2012). The amplitude of the P1 was defined by averaging ERP amplitudes during the time window between 130 ms and 170 ms after target onset on six occipital electrodes (Oz, O1, O2, POz, PO3, and PO4). These electrodes were centered over the
maximum of the ERP amplitudes at 150 ms after target onset as determined through visual
inspection of the scalp topography (see Fig. 2C).

A repeated measures ANOVA revealed that there was no significant effect of target
orientation ($F_{1,12} = 0.016, p = 0.901$) or interaction between target orientation and flanker
condition ($F_{1,12} = 0.002, p = 0.964$). We found a similar pattern of results—no significant effect
of target orientation or interaction—across all our ERP experiments. However, P1 amplitude
was suppressed in the “same” condition compared to the “orthogonal” condition ($F_{1,12} = 16.787,$
$p = 0.001$), consistent with orientation-specific surround suppression. Separate analyses for
each target orientation confirmed that the trend was similar for both vertical ($F_{1,12} = 7.654, p =
0.017$) and horizontal ($F_{1,12} = 13.455, p = 0.003$) targets (Fig. 4A).

After establishing orientation-specific surround suppression in the P1 amplitude of our
ERP data, we tested our grouping hypothesis using the same stimulus manipulations used in
Experiment 1. First in Experiment 2B, we increased the distance between the target and
flankers by doubling the center-to-center distance (6°, Fig. 4B) used in Experiment 2A. We
found no difference in P1 amplitude between the “same” and “orthogonal” conditions for both
target orientations (Fig. 4B; vertical, $F_{1,14} = 0.013, p = 0.911$; horizontal, $F_{1,14} = 0.081, p = 0.781$).

In Experiment 2C we used the distant-flanker configuration of Experiment 2B but made
the target and flankers appear to be grouped on a common surface that was distinct from the
background. Note that both the flankers and the cast-shadow surface appeared before the
onset of the target—using the same timing structure as Experiments 2A and 2B. Again,
consistent with our grouping hypothesis, P1 amplitude was suppressed in the “same” condition
compared to the “orthogonal” condition (Fig. 4C; vertical target, $F_{1,14} = 5.223, p = 0.038$;
horizontal target, $F_{1,14} = 4.461, p = 0.053$).

In Experiment 2D, we used the same spatial parameters of Experiment 2A—where we
observed strong orientation-specific surround suppression—but moved the flankers to different
surfaces than the target (Fig. 4D). Although the same distance manipulation resulted in strong
orientation-specific surround suppression in Experiment 2A (Fig. 4A), we found no evidence of
orientation-specific surround suppression in the P1 amplitude (Fig. 4D; vertical target, $F_{1,14} =
0.002, p = 0.965$; horizontal target, $F_{1,14} = 2.143, p = 0.165$) when the target and flankers were
placed on different surfaces.

Possible low-level, stimulus-based effects of the rendered surfaces on the target
response were minimized in Experiments 2C and 2D by having the surfaces appear first, before
target onset and remain until well after target offset. Note that we measured the visually
evoked potentials to target onset after the visually evoked potentials to flanker/shadow onset
diminished. However, there are still important stimulus-based differences to consider. In
particular there are horizontal edges in the proximity of the target (introduced by the surface)
in Experiment 2D that dodoes not exist in the other conditions. Can the lack of surround
suppression be explained by low-level stimulus-based differences? If the horizontal surface
edge behaved like a “flanker” it would have caused less suppression with a vertical target and
more suppression with a horizontal target. However, no such differences were observed in our
data (see Fig. 4D middle panel and compare the responses to the vertical target vs. the
horizontal target in the each condition).

To further rule-out possible low-level effects of the surfaces on the target response we
examined the “single” (target-only) condition in Experiment 2D: when the target was presented
without oriented flankers. If the surface was somehow interacting with the target stimulus in a low-level manner, a difference between the vertical and horizontal targets would be expected. Specifically, horizontal surface edges near the target in Experiment 2D should have resulted in little or no surround suppression for the vertical target if they acted like horizontally oriented flankers (e.g., like the vertical-target orthogonal condition in Experiment 2A). On the other hand, horizontal surface edges should have resulted in a suppressed response to the horizontal target (e.g., like the horizontal-target same condition in Experiment 2A). However, we found no significant difference between the vertical and horizontal targets when presented alone with the surfaces Experiment 2D (Fig. 5; \( t_{14} = 0.12, p = 0.91 \)) further suggesting that the surfaces do not have a simple, low-level effect on the visually evoked potentials to target onset.

We have demonstrated that grouping between the target and flankers was required for orientation-specific surround suppression in early visual cortex by showing that orientation-specific surround suppression occurred in experiments where the target could be grouped with the flankers (Experiments 2A and 2C) and that orientation-specific surround suppression did not occur in experiments where the target did not group with the flankers (Experiment 2B and Experiment 2D). To strengthen our claim, we tested whether there were significant interactions between stimulus conditions (same/orthogonal – a within-subjects factor) and surface manipulations for a given target-flanker distance (grouped/ungrouped – a between-subjects factor). In both the near flanker condition (Experiments 2A and 2D) and far flanker condition (Experiment 2B and 2C), there was a significant interaction between the stimulus condition and the surface manipulation (\( F_{1,26} = 9.454, p = 0.005 \) and \( F_{1,28} = 5.596, p = 0.025 \), respectively). These results suggest that our findings are not simply due to different statistical power in the individual experiments.

**Discussion**

Our results demonstrate that grouping—specifically mediated by the surface placement of the target and flankers—modulates orientation-specific surround suppression. We assume that the surface structure of the images in our experiments is represented in higher stages of the visual system that have neurons with sufficiently large receptive fields and complex tuning properties sensitive to relative depth. This may include regions such as the lateral occipital complex (Kourtzi and Kanwisher, 2001; Murray et al., 2002; 2003). Thus it is possible that feedback from these regions modulates orientation-specific surround suppression in early visual areas (e.g., V2-V3). However, the signal that we measured in response to the target—the P1 component—is believed to represent early, feedforward neural activity (Luck et al., 2000). Indeed, the onset of the P1 in our experiments (approximately 90-95 ms)—while likely too late to originate from V1—closely corresponds to the median onset time (approximately 85 ms) of single unit responses in V2 of the macaque monkey (Schmolesky et al., 1998). Thus, how do we reconcile the potential role of feedback with the modulation of an early feedforward neural signal? It is important to emphasize the relative timing of our stimulus presentation. The flankers and the cast-shadow surfaces (in Experiments 2C and 2D) were presented first, 1-2 seconds in advance of the briefly presented target stimulus. Thus, there was sufficient time for the putative feedback process to be in place and stabilized before the onset of the target. How our results generalize to other timing configurations—such as the simultaneous presentation of the target and flankers—remains an open question.
In Experiments 2C and 2D we used surfaces defined by cast-shadows to perceptually group or ungroup the target and flankers, respectively. It is likely that other grouping cues would serve a similar function and result in a similar modulation of orientation-specific surround suppression. For example, enclosing the target and flankers in Experiment 2B using lines could serve as a perceptual grouping cue (Palmer, 1999). Likewise, we predict that using binocular disparity manipulations to place the target on the same versus different depth plane as the flankers (Nakayama et al., 1989) would also lead to the presence versus absence of orientation-specific surround suppression.

We included a demanding central fixation task in Experiment 1 to specifically eliminate any differential effects of attention between the stimulus conditions. Further, in Experiment 2 the target stimuli were behaviorally irrelevant (i.e., under no specific task instruction), briefly flashed, of unpredictable orientation and peripherally located. Thus, it is unlikely that there are any simple confounds related to attention or motivation that could potentially explain our results.

Our stimulus configurations resemble those used to study crowding effects in the periphery where the sensitivity to a target is reduced when the target is surrounded by flankers (Bouma, 1970). Indeed, grouping also plays a critical role in crowding (Manassi et al., 2012). Despite some similarity in the stimulus configurations, our results were not due to crowding. First, the stimulus configurations were displayed near the fovea (displaced horizontally from the fixation point by 6° in Experiment 1 and 3° in Experiment 2) and perceptually the target was clearly visible. Second, we found that orientation-specific surround suppression was modulated by introducing surface structure even in the far-distance condition where the flankers were far removed from the target (6°). Vickery et al. (2009) showed that far-removed flankers that are outside of traditional crowding area could reduce the ability to identify a target when the target was masked. This result suggests that the mask that also affects target visibility interacts with the flankers in a super-additive way. In our experiments we did not manipulate the target visibility and simply measured the contrast detection threshold for an isolated target after adaptation (Experiment 1) and VEP for a briefly flashed target (Experiment 2). Thus, it is unlikely that our results were due to reduced target visibility by crowding. However, the question of whether the surface structure manipulations used in our experiments also affect crowding would be an interesting follow-up.

Overall, the high-level surface structure of the image—specifically, whether the target and flankers shared a common surface and thus were grouped into a single array—offers the most consistent explanation for our results. Indeed, our results are consistent with behavioral evidence that demonstrate a fundamental role of surface structure in perceptual grouping (Nakayama et al., 1989; Nakayama and Mackeben, 1989; Nakayama and Shimojo, 1992) and visual detection sensitivity (Huang et al., 2012). The results of the present study, together with our recent findings (Joo et al., 2012), suggest a coding scheme in early visual cortex that is sensitive to high-level image structure.

**Acknowledgements**

This research was supported by a National Science Foundation CAREER award (S.O.M.). We thank Alex Huk for his helpful comments on our manuscript.
References


Figure Legends
Fig. 1. The luminance profile of surface stimuli and the procedure of Experiment 1. (A) The surface stimuli used in Experiments 1C and 2C. (B) The surface stimuli used in Experiments 1D and 2D. X- and Y-axis represents the distance in visual angle from the center of the stimuli. The insets represent horizontal and vertical cross-sectional luminance profiles at the center of the stimuli. The color bars in each figure represent the gray scale luminance values (cd/m²) and they correspond to the values of y-axis in insets. (C) The procedure used in the behavioral contrast adaptation experiments. The “orthogonal” condition is shown. Observers initially adapted to the adapting pattern for 30 s before the first trial. A top-up adaptation period was inserted between the trials to maintain the adaptation state. A 0.5 s blank period was inserted between the adaptation period and the trial. During this blank period, the fixation point changed its luminance from black to white to cue the start of the trial.

Fig. 2. The general experimental conditions and ERP experimental procedure. (A) The stimulus configuration consisted of three Gabor stimuli: a central target with two flankers (one above and one below the target). The target orientation could be either vertical or horizontal. Flanker orientation matched the target orientation in the “same” condition (left) or was orthogonal to the target orientation in the “orthogonal” condition (right). (B) On a given trial, flankers (or flankers and cast-shadow in Experiments 2C and 2D) appeared before target onset. After a random duration between 1 and 2 s, the target was briefly flashed (100 ms). After target offset, flankers (or flankers and cast-shadow) remained in the display for 500 ms to ensure that flanker
offset did not contaminate the evoked potential to target onset. (C) Scalp topographical maps of the evoked potential at 150 ms after target onset, referenced to the average of the left and right mastoids for the vertical (left) and horizontal (right) target collapsed across flanker conditions in Experiment 1. The color bar represents voltage (µV). Each dot represents the 64 electrode recording sites. We averaged the waveforms across six occipital recording sites depicted by thicker dots to define the ERP signal.

Fig. 3. Stimulus configurations and the threshold ratio in each stimulus configuration in Experiment 1. (A) Near flankers. The center-to-center distance between the target and flankers was 3° and the target was displace by 6° horizontally from the fixation point. (B) Far flankers. The center-to-center distance between the target and flankers was 6°. (C) Far flankers on the same surface as the target. The center-to-center distance between the target and flankers was 6°. (D) Near flankers on a different surface from the target. The center-to-center distance between the target and flankers was 3°. * p < 0.05; ns, not significant. Error bars represent SEM across observers.

Fig. 4. The stimulus configurations and results of ERP Experiments. The left column shows the stimulus configurations in each experiment. The middle column shows P1 amplitude measured in each condition for each target orientation. Grey and black bars represent P1 amplitudes for the vertical target and horizontal target, respectively. The right column shows the ERP waveforms averaged across target orientations (vertical and horizontal). Dashed and solid lines represent the ERPs of the “orthogonal” and “same” condition, respectively. The shaded areas indicate the P1 amplitude measurement windows. (A) Experiment 2A: near flankers. The center-to-center distance between the target and flankers was 3°. (B) Experiment 2B: far flankers. The center-to-center distance between the target and flankers was 6°. (C) Experiment 2C: far flankers on the same surface as the target. The center-to-center distance between the target and flankers was 6°. (D) Experiment 2D: near flankers on a different surface from the target. The center-to-center distance between the target and flankers was 3°. * p < 0.05; ** p < 0.01; ns, not significant. Error bars represent within-subject 95% confidence interval (Loftus and Masson, 1994).

Fig. 5. The response to the single condition in Experiment 2D. The grey and black bars represent P1 amplitude for the vertical and horizontal targets, respectively. ns, not significant. Error bars represent within-subject 95% confidence interval (Loftus and Masson, 1994).