Activity in visual area V4 correlates with surface perception

Seth E. Bouvier   Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA

Kristen S. Cardinal   Department of Neurology, University of California, Los Angeles, CA, USA

Stephen A. Engel   Department of Psychology, University of Minnesota, Minneapolis, MN, USA

The neural mechanisms responsible for unifying noncontiguous regions of a visual image into a percept of a single surface remain largely unknown. To investigate these mechanisms, we used a novel stimulus in which local luminance was the only cue for surface segmentation. Subjects viewed an array of small adjoining elements that were randomly assigned as either surface or noise every 100 ms. On each trial, the luminance of surface elements was fixed to a single value and the luminance of noise elements was randomly assigned. As the ratio of surface to noise elements changed, subjects perceived either a surface embedded in noise or noise alone. In three functional magnetic resonance imaging (fMRI) experiments, early visual area V1 responded most strongly during trials with a low surface-to-noise ratio while later areas responded most strongly during trials with a high ratio. Furthermore, even at identical surface-to-noise ratios, responses in area V4 were higher during trials in which the subject perceived a surface than during trials in which the subject did not. Early visual areas did not show this pattern. These results suggest that visual area V4 contains neurons critical for the representation of surfaces.

Introduction

Many psychological principles governing image segmentation have been described (reviewed by Rock & Palmer, 1990), yet little is known about the neural mechanisms that underly this process. A critical step in segmentation process is the assignment of regions of the visual image to continuous surfaces. There is evidence that such surface representations are themselves a fundamental unit of perception (reviewed by Nakayama, He, & Shimojo, 1995). Inferring the presence of a surface is a difficult problem because parts of the surface may be occluded, causing it to be fragmented in the image.

The visual system uses a variety of cues to infer the presence of continuous surfaces. Useful cues include stereo depth, contour continuity, motion, recognizable form, and others (reviewed by Albright & Stoner, 2002; Nakayama et al., 1995; Northdurft, 1994; Roelfsema, 2006; Sasaki, 2007). The cues exploit some of the statistical regularities of natural scenes (reviewed by Geisler, 2008).

As one important example, visible surfaces tend to be relatively uniform in color. Analysis of image statistics has revealed that two pixels on the same surface will likely have the same luminance and color, and that two pixels on different surfaces will likely differ in both luminance and color (Ruderman, Cronin, & Chiao, 1998). Fine, MacLeod, and Boynton (2003) used the statistics of natural images as the priors for a Bayesian model that predicted whether nearby pixels in an image arose from the same underlying surface. Their model matched human judgments fairly well in a surface perception task, which suggests that humans do make use of color image statistics. The neural computations that identify surfaces from common color remain unknown, however.

The experiments reported here were designed to identify regions in cortex that might play a key role in such computations. As a first step, we used colors that were shades of gray differing in their luminance. We presented subjects with dynamic displays of elements that belonged to two overlapping distributions of luminance values (see Figure 1). One distribution simulated a uniform surface, and was very narrow in its luminance range. The second distribution was noise that spanned most of the range of displayable luminance values. Subjects often perceived the displays to contain a continuous uniform surface covered by discontinuous patches of texture. Subjects' perceptions varied with the statistical properties of the displays; ones containing a higher ratio of surface elements to noise elements were
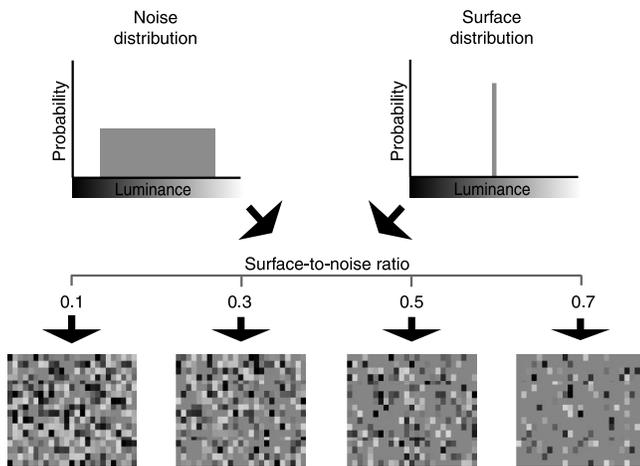
Figure 1. Stimulus Generation. The value of each element in the display was drawn from one of two probabilistic distributions of luminance values. In the noise distribution (top left), all luminance values except the darkest and lightest were equally likely, while in the surface distribution (top right) only a single luminance value was possible. At each location in the display, an element was randomly assigned to one of the two distributions according to the surface-to-noise ratio of the stimulus, and its luminance value determined accordingly. Shown in the bottom row are frames created at the surface-to-noise ratios of 0.1, 0.3, 0.5, and 0.7, from left to right. A new frame was presented every 100 ms, and the luminance of each element was reassigned following this same procedure.

perceived more often as containing surfaces. Using functional MRI, we identified a region in extra-striate cortex where neural activity correlated with perception. This area is a likely substrate for neural representations that underlie surface perception.

Methods

Subjects

Five subjects participated in experiments 1 and 4 (mean age 30.2 ± 2.7 years). Six subjects participated in experiments 2 and 3 (mean age 30 ± 2.2 years). Subjects were all right-handed and had normal or corrected-to-normal vision. Procedures were approved by the UCLA Office for the Protection of Research Subjects.

Stimuli

We used a novel dynamic stimulus that consisted of a rectangular display (23×16.7 degrees of visual angle) divided into 59 by 44 small square elements, each approximately 0.38 degrees in size (see Figure 1). The

experiments reported here used only grayscale colors; all elements were the mean chromaticity of the monitor, but differed in their luminance.

For each frame of the display, the luminance of every element was drawn from one of two distributions of luminance values. The first distribution (called the noise distribution) contained a range of luminance values, and the second distribution (called the surface distribution) contained only one possible luminance value. The noise distribution was a uniform distribution that spanned the range from 20% to 80% of the maximum luminance the display could produce. This distribution did not change from trial to trial. The surface luminance value varied from trial to trial, and fell within a constrained range between the brightest and darkest 33% of noise luminance values.

Each element in the frame was assigned to one of the distributions, with a probability that we termed the "surface-to-noise ratio" of the stimulus. For example, a ratio of 1 indicated that all the luminance values were chosen from the surface distribution; 0 meant that all the luminance values were chosen from the noise distribution, and 0.5 meant that on average half the values were chosen from each distribution. On successive frames of a stimulus, each element was independently reassigned, and the surface-to-noise ratio remained constant throughout the duration of the stimulus. Every 100 ms, a new frame was presented, in which the elements were a new random draw from the two distributions. Examples of single frames of the stimulus at different surface-to-noise ratios are shown in Figure 1.

Because small fixation marks disrupted surface perception, we used a thin cross spanning the entire stimulus as a fixation mark. This created the appearance of a window frame through which the stimulus was viewed.

Procedure

We conducted four experiments using the dynamic stimuli. The experiments differed slightly in procedure (see Figure 2). Experiment 1 measured behavior only, and subjects reported if they perceived a surface in displays of different surface-to-noise ratios. Experiment 2 measured neural activity with fMRI while subjects passively viewed stimuli with high and low surface-to-noise ratio displays. Experiments 3 and 4 combined behavior with fMRI. Subjects reported if they perceived a surface in the displays and also performed an attention-controlling task. In experiment 3, subjects viewed stimuli with fixed surface-to-noise ratios of 0.1, 0.3, 0.5, and 0.7 and in experiment 4, subjects viewed stimuli with signal-to-noise ratios at, above, and below a threshold value determined in experiment 1. Experiments were presented with MR-compatible goggles (Resonance Technologies, Inc.) controlled by computers programmed in MATLAB using the Psychophysics toolbox routines (Brainard, 1997).

Experiment 1

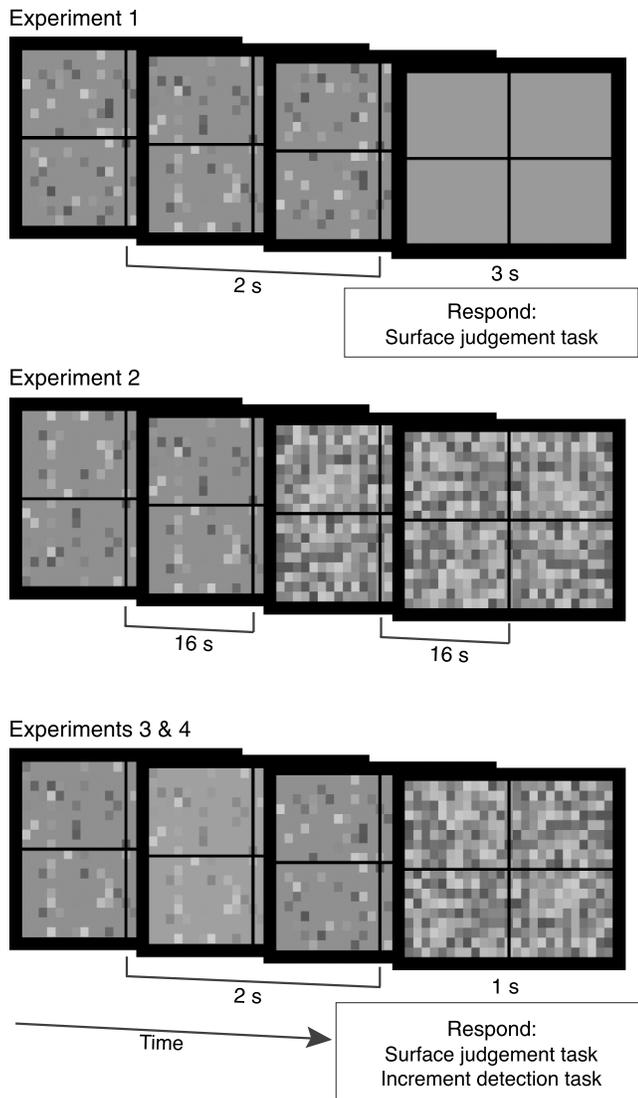**Experiment 2**

Experiments 3 & 4

Figure 2. Stimuli and Tasks. In experiment 1, subjects viewed the dynamic surface display for two seconds followed by three seconds of a gray screen during which they rate the stimulus as appearing to contain a surface embedded in noise or noise alone. In experiment 2, the display alternated between sixteen-second blocks of high and low surface-to-noise ratio stimuli. Subjects performed no task. In experiments 3 and 4, subjects viewed two seconds of the dynamic surface display followed by a one second 100% static noise response period. Subjects performed the Surface Judgment Task, as in experiment 1, and also reported small changes in the luminance of the surface display, when present, by a second button press. The second frame, which has a higher mean luminance than the trial's other frames, illustrates an increment target. In experiment 3, subjects viewed fixed surface-to-noise ratios of 0.1, 0.3, 0.5, and 0.7; in experiment 4, subjects viewed surface-to-noise ratios that depended on their individual threshold.

Experiment 1

A tone indicated the start of a trial and was followed by two seconds of the dynamic stimulus. Subjects then performed a two alternative forced choice *surface judgment task,* in which they reported by key press their percept as either a surface embedded in noise, or noise alone. For three seconds after each stimulus, the screen was a uniform gray while subjects made their responses.

From trial to trial, both the surface-to-noise ratio and the luminance value of the surface varied under experimental control. The surface-to-noise ratio was adjusted using a staircase procedure and the surface luminance was randomly selected on each trial. Subjects completed a single session of 243 trials.

Experiment 2

Subjects passively viewed the displays while fMRI data were acquired. The dynamic display alternated every 16 seconds between a surface-to-noise ratio of 0.9 and a surface-to-noise ratio of 0.1. In total, there were twelve cycles of high and low surface-to-noise ratios. Subjects participated in a single fMRI scan only.

Experiment 3

Subjects viewed a sequence of many short trials in a rapid, event-related fMRI design. A tone indicated the start of each trial during which subjects viewed two seconds of the dynamic stimulus. Subjects performed the surface judgment task and responded during a one second interval, during which the display was composed of 100% static noise. This was meant to inhibit the perception of a uniform surface between trials.

Subjects also performed a second, attention-controlling task, the *increment detection task.* During 50% of the trials, the luminance value of the surface distribution changed for a single 100 ms frame of the dynamic display. The new luminance value of the surface was either slightly lighter or darker than the original surface, and the magnitude of the shift was controlled by a staircase procedure. Separate, independent staircases were used for each surface-to-noise ratio. Subjects reported the direction of these brief luminance shifts in addition to performing the surface judgment task. The shifts occurred at random times constrained to be between 0.8 and 1.7 seconds after the start of the stimulus. Subjects made their responses using a MR compatible button box.

In each trial, the dynamic display was set to one of four possible surface-to-noise ratios: 0.1, 0.3, 0.5, and 0.7. Additional "null" trials containing the fixation mark and a static 100% noise display were presented as a baseline for data analysis. The four surface-to-noise ratios (and null trials) were presented in an unpredictable, but counterbalanced, order using an m-sequence (Buracas & Boynton, 2002). Each fMRI scan contained 125 trials, and subjects participated in 3–4 scans per experiment.

Experiment 4

The methods used in experiment 4 were the same as in experiment 3, with the following exceptions: We presented subjects with stimuli at their threshold surface-to-noise ratio calculated from data gathered in experiment 1. To prevent subjects from noticing the repeated presentation of threshold level trials and basing their responses on memory of past trials, we also included surface-to-noise ratios 0.25 above and below their threshold levels. In addition to the three threshold-dependent noise levels, static noise null trials were presented as in experiment 3. Each fMRI scan had 128 trials.

MR data acquisition

Subjects participated in multiple scanning sessions, during which brain activation was measured using blood oxygenation level dependent, or BOLD, fMRI (Kwong et al., 1992; Ogawa et al., 1992). In the first scanning session, subjects viewed standard retinotopic stimuli to identify visual areas (DeYoe et al., 1996; Engel et al., 1994; Sereno et al., 1995). In addition, two high-resolution T1 weighted anatomical scans (MPRAGE) were acquired for use in tissue segmentation and unfolding algorithms. Functional data using the surface stimuli were collected in a subsequent session. For all functional runs, twelve slices of fMRI data, oriented perpendicular to the calcarine fissure, were acquired using an EPI sequence (TR = 1000 ms; TE = 45 ms; voxel size = 3.1 × 3.1 × 4 mm). All brain imaging was conducted using the 3-Tesla Siemens Allegra MR scanner located at the UCLA Ahmanson-Lovelace Brain Mapping Center in Los Angeles, CA.

Analysis

Behavior

For each subject, the percent of "surface" responses in the surface judgment task was calculated for each surface-to-noise ratio presented. A Weibull function was fit to these data to estimate a psychometric function. Each subject's threshold level was calculated as the value of the Weibull function at 50% surface responses.

fMRI

Flattened cortical maps were generated from the MPRAGE scans using SurfRelax (Larsson, 2001). Data from the retinotopy scans were projected onto the flat maps using mrVista (http://white.stanford.edu/software/). Visual areas were identified using reversals in phase-encoded polar angle retinotopy data (DeYoe et al., 1996; Engel et al., 1994; Sereno et al., 1995). Retinotopic regions of interest (ROIs) were drawn by hand, based on the polar angle reversals as they appeared on the flat map projections.

Motion during each scan was corrected using an intensity-based linear method (Jenkinson, Bannister, Brady, & Smith, 2002). To correct for any head motion between scans, functional run were coregistered using a rigid body transformation (Jenkinson et al., 2002).

The raw fMRI time course from each voxel was converted to percent change by subtracting and dividing by the voxel's mean activity during the scan. Next, mean timecourses were computed by averaging the activity in all the voxels in each region of interest. Because the size of the surface stimuli was approximately equal to the size of the retinotopy stimuli, the entire bilateral retinotopic ROIs were included in the analysis, without statistical thresholding.

In the blocked design experiment (Experiment 2), we averaged each subject's twelve cycles together. The amplitude of the response was estimated by fitting a sinusoid to the average data from each region. The sinusoid was of fixed phase and peaked during the middle of the high surface-to-noise block. We used the best-fitting amplitude of the sinusoid as a measure of the region's response to the blocked scan.

To estimate responses for each condition in the rapid event-related design we fit a standard model (a finite impulse response model) to the fMRI data to produce an estimated hemodynamic response function (HRF) for each condition. The model contained one event for every stimulus presentation (trial), and each condition was modeled with one parameter per TR (1 second) for 16 seconds. The model was fit using ordinary least-squares. The amplitude of response in each condition was calculated as the peak (maximum) value of the estimated HRF.

To evaluate the parametric effect of surface-to-noise ratio on neural responses in experiment 3, a line was fit to the peak responses to the four surface-to-noise ratios, for each subject and in each visual area. The slope of this line was used as a measure of the effect of surface-to-noise ratio changes on the neural response, called here the Response Slope. A Response Slope of 1 indicates that increasing the surface-to-noise ratio by 0.1 increased the response in the visual area by 0.1% MR signal change, for example. Quality of the linear fits was evaluated using a correlation coefficient, and statistical significance of Response Slopes was tested with the Wilcoxon signed rank test for zero median.

In experiment 4, we extracted HRFs for each noise level separately, but further analyzed only the responses to the threshold level trials. To evaluate the relationship between percept and neural response, we subtracted the peak of the response during trials that led to a surface percept from the peak of the response during trials that led to a noise percept. The statistical significance of this difference was tested using a one-tailed version of the (paired) Wilcoxon rank sum test for equal medians.

Stimuli

Because the fMRI data in experiment 4 were analyzed according to the subjects' percept, it was important to be sure that the stimuli perceived as containing a surface did not differ from those that were perceived as noise alone. To test for such differences we converted each frame of all stimuli presented into contrast by subtracting and dividing by the image's mean. We then calculated a spatial frequency spectrum for each image by summing its Fourier transform amplitude across polar angle. The average spatial frequency content of trials rated as containing surfaces was compared with those that were rated as noise alone. As a measure of total image contrast, we computed the variance of the contrast of all pixels in the image, and compared this quantity from trials rated as containing a surface to trials that were rated as noise alone. This variance is the square of RMS contrast.

Results

Experiment 1

Subjects reported that at low surface-to-noise ratios, the stimulus appeared to be primarily dynamic noise while, at high surface-to-noise ratios, the stimulus appeared to be a surface partially occluded by dynamic noise. Subjects performed the surface judgment task (see Methods) to determine their individual threshold psychometric functions; results are shown in Figure 3. In all cases, the percentage of surface responses increased systematically with an increase in surface-to-noise ratio.

For use in Experiment 4, we calculated each subject's threshold—the surface-to-noise ratio equally likely to evoke a percept of a surface as a percept of noise alone. A Weibull function was fit to each subject's behavioral data and the threshold was estimated as the surface-to-noise ratio corresponding to 50% surface responses. The mean threshold was 0.72 ± 0.05 (standard error).

Experiment 2

Subjects viewed the dynamic surface stimulus, alternating between high (0.9) and low (0.1) surface-to-noise ratios. Activity in calcarine regions tended to be out of phase with activity in more ventral locations (Figure 4A). Specifically, in V1, responses were greater during the low surface-to-noise ratio blocks, while in V4, responses were greater during the high surface-to-noise ratio blocks. To quantify these effects, we fit the activity in each subject's visual areas with a sinusoid that peaked during the middle of the high surface-to-noise ratio block. Figure 4C plots the averaged, best-fitting sinusoid amplitudes. The amplitude is most negative in area V1, most positive in area V4, and intermediate in the other visual areas.

Experiment 3

Subjects viewed stimuli with surface-to-noise ratios of 0.1, 0.3, 0.5, and 0.7, and performed both the surface judgment and increment detection tasks in the scanner. Subjects correctly performed the increment detection task in 77.6% of trials overall, 77.9% when a surface was perceived and 77.4% when no surface was perceived. These results suggest that visual attention was equally engaged during both trial types.

In V1, the response strength increased with decreasing surface-to-noise ratios, while in V4 the opposite pattern was observed; the response strength increased with increasing surface-to-noise ratio. In Figure 5, the responses
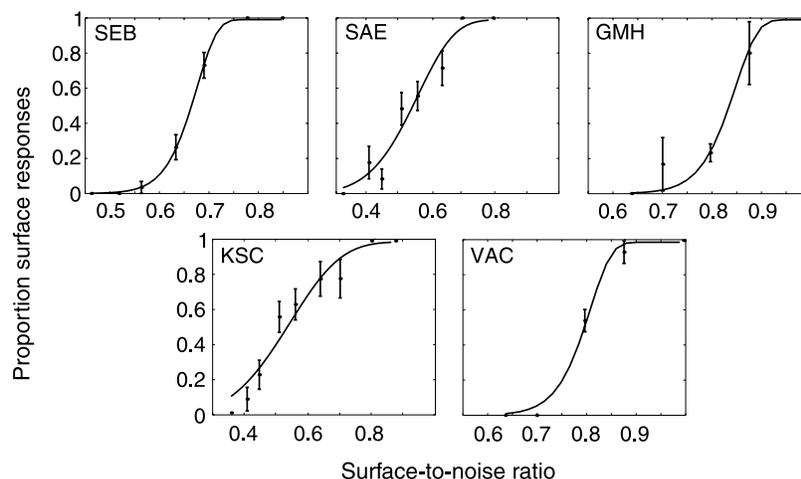

Figure 3. Experiment 1 results. Shown are the results of the behavioral experiment for all five subjects. The line is a Weibull function fit to the data. In all figures, error bars are ± one *SEM*.
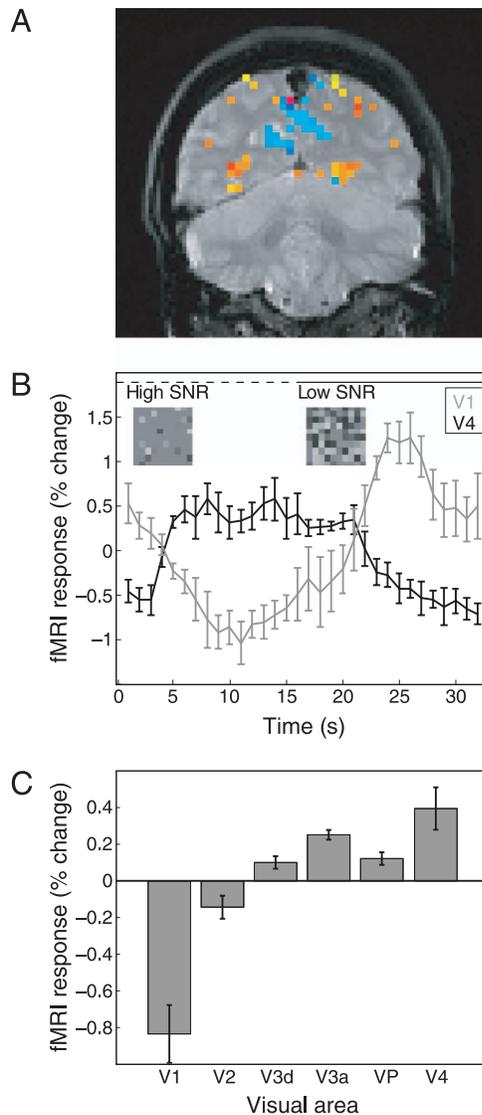
Figure 4. Experiment 2 results. (A) Shown is a coronal slice taken from one subject in grayscale. Overlaid, in color, is the correlation between the subject's fMRI activity during the scan and a sinusoid with a period of 32 seconds. The hot colors indicate a positive correlation and the cool colors indicate a negative correlation. Most of the blue pixels in the slice fell within visual area V1, and most of the orange pixels fell within visual area V4. (B) Plotted are the average responses for this subject, of visual areas V1 (gray) and V4 (black) to alternating blocks of high and low surface-to-noise stimuli. The high surface-to-noise stimulus appeared for the first 16 seconds, and the low surface-to-noise stimulus for the final 16 seconds of each cycle. (C) Cross-subject average amplitude of responses in visual areas V1–V4.

from areas V1 and V4 are plotted for each surface-to-noise ratio.

To quantify the relationship between surface-to-noise ratio and response, a line was fit to the peak responses to the four surface-to-noise ratios. The slope of this line was termed the Response Slope, and it was measured for each

visual area. Only V1 and V4 showed Response Slopes significantly different from zero ($p < .05$). In V1, the Response Slope was negative, indicating that increasing the surface-to-noise ratio decreased the response size. In V4, the Response Slope was positive, indicating that increasing the surface-to-noise ratio increased the response size. Areas V3d, VP, and V3a showed nonsignificant trends for positive slope.

Experiment 4

Subjects viewed stimuli at three levels during fMRI scans: their individual surface-to-noise threshold level as determined in experiment 1, a value 0.25 above threshold
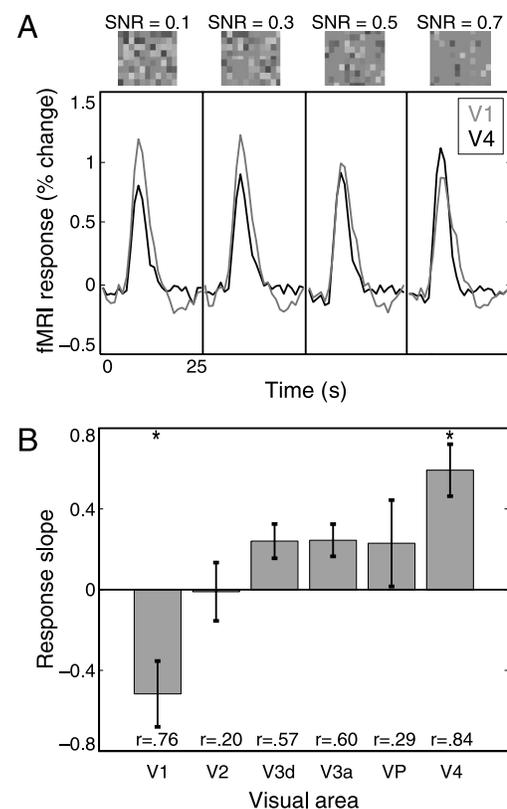

Figure 5. Experiment 3 results: Effects of surface-to-noise ratio on visual area responses. (A) Plotted are the responses of visual areas V1 (gray) and V4 (black) to stimuli with surface-to-noise ratios of 0.1, 0.3, 0.5, and 0.7, from left to right. In V1, the responses increase with decreasing surface-to-noise ratios ($p < .05$), and in V4 the responses increase with increasing surface-to-noise ratios ($p < .05$). (B) Shown, for each visual area, is the average Response Slope. A slope of 0.5, for example, indicates that a change in surface-to-noise ratio of 0.2 produced a 0.1% change in fMRI response amplitude in that visual area (for details see Methods). Only visual areas V1 and V4 showed Response Slopes that were significantly different than zero (*; $p < .05$). Average correlation coefficients between the line used to compute Response Slope and fMRI response amplitudes are shown for each visual area.

and a value 0.25 below threshold. On average, subjects perceived a surface in 41% of the threshold noise level trials, in 98% of trials with a noise level 0.25 above threshold and in 5% of trials with a noise level 0.25 below threshold. Performance on the increment detection task was 71.5% overall (at all three surface-to-noise levels). On threshold-level trials, the subjects responded correctly on 75.8% of trials in which they perceived a surface and 72.8% of trials in which they did not perceive a surface. This small difference was not statistically reliable.

Analysis of the fMRI data was restricted to the threshold-level trials. Trials were sorted into surface and noise trials based upon subjects' reported percepts in the surface judgment task, and fMRI responses were calculated separately for each trial type. The results from visual areas V1 and V4 are plotted in Figure 6A. In V1, the responses during the surface and the noise trials did not differ reliably, but the two responses did differ in V4 ($p < .05$).

The differences between the peak responses during the surface and the noise trials are shown in Figure 6B. Only
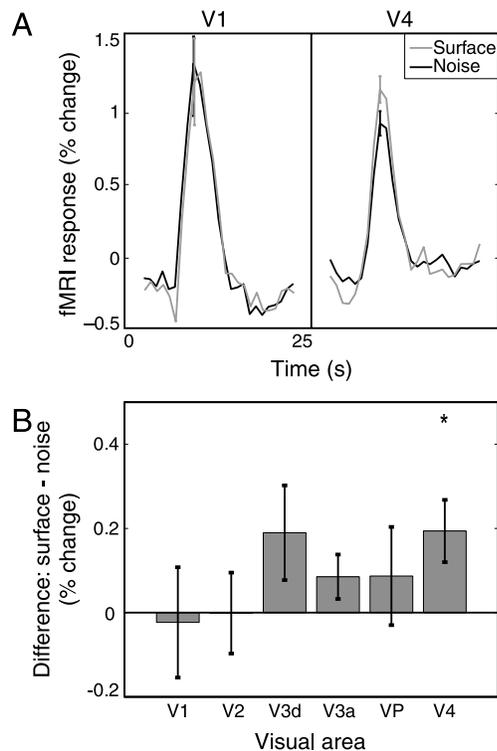

Figure 6. Experiment 4 results: Effect of surface perception on visual area responses. (A) The responses to individual threshold level surface-to-noise stimuli are shown for visual area V1 (left panel) and V4 (right panel). In black are the responses during the trials that led to the percept of noise alone, and in gray are the responses during the trials that led to the percept of a surface. The difference between the peak height was significant in area V4 ($p < .05$), but not in area V1. (B) Shown for each visual area is the difference between the response peaks during trials in which a surface was perceived and trials in which no surface was perceived. Only visual area V4 showed a difference significantly different than zero (*; $p < .05$).

visual area V4 had a difference significantly different from zero, though areas V3d, VP, and V3a showed non-significant positive trends.

Could aspects of the stimulus explain the differences between visual areas seen in Experiments 2, 3, and 4? Neurons in V4, for example, are known to prefer lower spatial frequencies than neurons in V1. Because the luminance of each element in the display is an independent random variable, however, there should be no spatial or temporal correlations in the images used here (Chubb, Econopouly, & Landy, 1994; Victor, Chubb, & Conte, 2005). That is, the spatial frequency spectra of our stimuli should be flat, and identical, at all surface-to-noise ratios in all our experiments. Accordingly, differential sensitivity to spatial frequency cannot account for the results of experiments 2 and 3. However, subjects in experiment 4 may have based their judgments on small random variations in stimulus contrast or spatial frequency content of the stimuli.

To test whether this was the case, we compared the average total contrast and average spatial frequency spectra of the stimuli judged to contain a surface with those that were judged to be noise alone. The distribution of energy was flat across spatial frequencies, for both trial types and did not differ significantly for any spatial frequency band. The average difference in energy across all spatial frequency bands was very small (less than 0.1% of the total energy) and did not differ significantly between trial types. Finally, the average variance of stimulus contrast within an image was not significantly different between the two conditions. The average difference across subjects was less than 0.02% of the total variance. Thus it is unlikely that differences in stimulus contrast or spatial frequency content affected the results of experiment 4.

Discussion

We used a novel stimulus to study the neural basis of surface perception. Subjects viewed dynamic displays whose elements were drawn from overlapping probabilistic distributions of luminance values. Under certain conditions, the elements belonging to one distribution appeared to belong to a single continuous surface. Changing the proportion of elements drawn from the two distributions altered this perception of a unified surface; as the ratio of surface elements to noise elements decreased, subjects were less likely to report having perceived a surface during the trial.

Previous studies have demonstrated that subjects are sensitive to related statistical properties of artificial stimuli (Chubb et al., 1994; Goda & Fujii, 2001; Li & Lennie, 1997, 2001), and of the natural world (Fine et al., 2003). However, the present study uses the statistical properties

of the two distributions to evoke the percept of the segmentation of a uniform surface in a noisy environment. No prior studies have manipulated luminance image statistics alone to evoke a surface percept.

More importantly, we identified visual areas that also showed an effect of stimulus surface-to-noise ratio. In experiments 2 and 3, responses in early area V1 decreased with surface-to-noise ratio, while responses in later visual areas increased. The increase in activity during high noise trials seen in V1 is likely explained by those trials' greater amounts of local spatiotemporal contrast energy. The increased activity in later visual areas during the low noise trials is likely due to the perception of a unified surface during those trials.

Further support for this explanation comes from experiment 4, which controlled for stimulus differences. Brain activity was compared during trials that had identical surface-to-noise ratios, but resulted in different percepts. We found that responses in V4 were higher when a surface was perceived and lower when noise alone was perceived, even when the surface-to-noise ratio was held constant. This pattern was not observed in other visual areas.

Because experiment 4 attempted to equate the contrast and spatial frequency content of the stimuli, it was important to verify that the stimuli during trials rated as containing a surface were not different than stimuli during trials rated as noise alone. We examined the spatial frequency content and total contrast of all the stimuli presented to subjects during the threshold trials and found no differences between the trials rated as containing a surface and trials rated as noise alone.

Since area V4 is known to be sensitive to attentional manipulations, an alternative explanation of the correlation between activity in V4 and surface perception is that the percept of a surface either causes or is caused by an increase in attention on that trial. There are documented examples of interactions between attention and segmentation (reviewed by Driver, Davis, Russell, Turatto, & Freeman, 2001). On the other hand, figure/ground enhancement can be largely independent from attention (Marcus & Van Essen, 2002). Here, we controlled for attention by having subjects engaged in all trials in a challenging luminance increment detection task. Thus, it is unlikely that subjects' attentional state varied between conditions. Furthermore, area V1, which is also known to be sensitive to attentional manipulations, showed opposite patterns from V4 in experiments 2 and 3, and did not show an effect of surface perception in experiment 4.

A prior report found that successful perceptual organization reduced activity in V1 (Murray, Kersten, Olshausen, Schrater, & Woods, 2002). Similar effect were not seen in experiment 4, most probably because even when our subjects reported seeing an organized surface they also perceived large amounts of unorganized noise accompanying it in the display.

The computations underlying scene segmentation have been presumed to occur at the earliest levels of the visual system, due in part to the importance of a segmented image to many visual processes (reviewed by Palmer, Brooks, & Nelson, 2003). Our results suggest that surface segmentation, at least based on grayscale color cues, is likely to depend on computations performed in intermediate visual areas. Our data most strongly link area V4 to these computations, though we cannot rule out roles for areas V3d, V3a, and VP. The delayed enhancement of neural responses related to segmentation observed in earlier visual areas (e.g., Hupé et al., 1998; Lamme, 1995) may be the result of recurrent connections between these areas and area V4.

Conclusions

We presented subjects with a simple stimulus that engaged surface perception mechanisms. Both subjects' frequency of seeing of a surface, and neural responses in area V4 increased as we increased the ratio of surface to noise elements in the stimulus. Moreover, during trials with a threshold ratio of surface and noise elements, activity in V4 was higher on trials in which a surface was seen, and lower on trials in which only noise was perceived. Earlier visual areas did not show these effects. Visual area V4 likely contains neurons that are important for encoding properties of visible surfaces.

Acknowledgments

Commercial relationships: none.
Corresponding author: Seth Bouvier.
Email: sbouvier@princeton.edu.
Address: Center for the Study of Brain, Mind and Behavior, Green Hall Princeton University, Princeton, NJ 08540 USA.

References

Albright, T. D., & Stoner, G. R. (2002). Contextual influences on visual processing. *Annual Review of Neuroscience, 25,* 339–379. [PubMed]

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision, 10,* 433–436. [PubMed]

Buracas, G. T., & Boynton, G. M. (2002). Efficient design of event-related fMRI experiments using M-sequences. *Neuroimage, 16,* 801–813. [PubMed]

Chubb, C., Econopouly, J., & Landy, M. S. (1994). Histogram contrast analysis and the visual segregation

of IID textures. *Journal of the Optical Society of America A, Optics, Image Science, and Vision, 11,* 2350–2374. [PubMed]

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

Driver, J., Davis, G., Russell, C., Turatto, M., & Freeman, E. (2001). Segmentation, attention and phenomenal visual objects. *Cognition, 80,* 61–95. [PubMed]

Engel, S. A., Rumelhart, D. E., Wandell, B. A., Lee, A. T., Glover, G. H., Chichilnisky, E. J., et al. (1994). fMRI of human visual cortex. *Nature, 369,* 525. [PubMed]

Fine, I., MacLeod, D. I., & Boynton, G. M. (2003). Surface segmentation based on the luminance and color statistics of natural scenes. *Journal of the Optical Society of America A, Optics, Image Science, and Vision, 20,* 1283–1291. [PubMed]

Geisler, W. S. (2008). Visual perception and the statistical properties of natural scenes. *Annual Review of Psychology, 59,* 167–192. [PubMed]

Goda, N., & Fujii, M. (2001). Sensitivity to modulation of color distribution in multicolored textures. *Vision Research, 41,* 2475–2485. [PubMed]

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. [PubMed]

Jenkinson, M., Bannister, P., Brady, M., & Smith, S. (2002). Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage, 17,* 825–841. [PubMed]

Kwong, K. K., Belliveau, J. W., Chesler, D. A., Goldberg, I. E., Weisskoff, R. M., Poncelet, B. P., et al. (1992). Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences of the United States of America, 89,* 5675–5679. [PubMed] [Article]

Lamme, V. A. (1995). The neurophysiology of figure-ground segregation in primary visual cortex. *Journal of Neuroscience, 15,* 1605–1615. [PubMed] [Article]

Larsson, J. (2001). Imaging vision: Functional mapping of intermediate visual processes in man. PhD thesis. Karolinska Institutet, Stockholm, Sweden.

Li, A., & Lennie, P. (1997). Mechanisms underlying segmentation of colored textures. *Vision Research, 37,* 83–97. [PubMed]

Li, A., & Lennie, P. (2001). Importance of color in the segmentation of variegated surfaces. *Journal of the Optical Society of America A, Optics, Image Science, and Vision, 18,* 1240–1251. [PubMed]

Marcus, D. S., & Van Essen, D. C. (2002). Scene segmentation and attention in primate cortical areas V1 and V2. *Journal of Neurophysiology, 88,* 2648–2658. [PubMed] [Article]

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 of the United States of America, 99,* 15164–15169. [PubMed] [Article]

Nakayama, K., He, Z. J., & Shimojo, S. (1995). Visual surface representation: A critical link between lower-level and higher-level vision. In S. M. Kosslyn & D. N. Osherson (Eds.), *Visual cognition and action.* Cambridge, MA: MIT.

Nothdurft, H. C. (1994). Common properties of visual segmentation. *Ciba Foundation Symposium, 184,* 245–259. [PubMed]

Ogawa, S., Tank, D. W., Menon, R., Ellermann, J. M., Kim, S. G., Merkle, H., et al. (1992). Intrinsic signal changes accompanying sensory stimulation: Functional brain mapping with magnetic resonance imaging. *Proceedings of the National Academy of Sciences of the United States of America, 89,* 5951–5955. [PubMed] [Article]

Palmer, S. E., Brooks, J. L., & Nelson, R. (2003). When does grouping happen? *Acta Psychologica, 114,* 311–330. [PubMed]

Peterson, M. A. (1994). Object recognition processes can and do operate before figure-ground organization. *Current Directions in Psychological Science, 3,* 105–111.

Rock, I., & Palmer, S. (1990). The legacy of Gestalt psychology. *Scientific American, 263,* 84–90. [PubMed]

Roelfsema, P. R. (2006). Cortical algorithms for perceptual grouping. *Annual Review of Neuroscience, 29,* 203–227. [PubMed]

Ruderman, D. R., Cronin, T. W., & Chiao, C. C. (1998). Statistics of cone responses to natural images: Implications for visual coding. *Journal of the Optical Society of America A, 15,* 2036–2045.

Sasaki, Y. (2007). Processing local signals into global patterns. *Current Opinion in Neurobiology, 17,* 132–139. [PubMed]

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

Victor, J. D., Chubb, C., & Conte, M. M. (2005). Interaction of luminance and higher-order statistics in texture discrimination. *Vision Research, 45,* 311–328. [PubMed]