Feedforward and Recurrent Processing in Scene Segmentation: Electroencephalography and Functional Magnetic Resonance Imaging

H. Steven Scholte[1], Jacob Jolij[1]*, Johannes J. Fahrenfort[1], and Victor A. F. Lamme[1,2]

Abstract

■ In texture segregation, an example of scene segmentation, we can discern two different processes: texture boundary detection and subsequent surface segregation [Lamme, V. A. F., Rodriguez-Rodriguez, V., & Spekreijse, H. Separate processing dynamics for texture elements, boundaries and surfaces in primary visual cortex of the macaque monkey. *Cerebral Cortex, 9,* 406–413, 1999]. Neural correlates of texture boundary detection have been found in monkey V1 [Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J., & Davis, J. Visual cortical mechanisms detecting focal orientation discontinuities. *Nature, 378,* 492–496, 1995; Grosof, D. H., Shapley, R. M., & Hawken, M. J. Macaque-V1 neurons can signal illusory contours. *Nature, 365,* 550–552, 1993], but whether surface segregation occurs in monkey V1 [Rossi, A. F., Desimone, R., & Ungerleider, L. G. Contextual modulation in primary visual cortex of macaques. *Journal of Neuroscience, 21,* 1698–1709, 2001; Lamme, V. A. F. The neurophysiology of figure ground segregation in primary visual-cortex. *Journal of Neuroscience, 15,* 1605–1615, 1995], and whether boundary detection or surface segregation signals can also be measured in human V1, is more controversial [Kastner, S., De Weerd, P., & Ungerleider, L. G. Texture seg-gregation in the human visual cortex: A functional MRI study. *Journal of Neurophysiology, 83,* 2453–2457, 2000]. Here we present electroencephalography (EEG) and functional magnetic resonance imaging data that have been recorded with a paradigm that makes it possible to differentiate between boundary detection and scene segmentation in humans. In this way, we were able to show with EEG that neural correlates of texture boundary detection are first present in the early visual cortex around 92 msec and then spread toward the parietal and temporal lobes. Correlates of surface segregation first appear in temporal areas (around 112 msec) and from there appear to spread to parietal, and back to occipital areas. After 208 msec, correlates of surface segregation and boundary detection also appear in more frontal areas. Blood oxygenation level-dependent magnetic resonance imaging results show correlates of boundary detection and surface segregation in all early visual areas including V1. We conclude that texture boundaries are detected in a feedforward fashion and are represented at increasing latencies in higher visual areas. Surface segregation, on the other hand, is represented in "reverse hierarchical" fashion and seems to arise from feedback signals toward early visual areas such as V1. ■

INTRODUCTION

Scene segmentation is fundamental to the process of perceptual organization. When scene segmentation is based on a difference in texture, the brain has to evaluate information from large parts of the visual field to group features into coherent objects and their surroundings. The top row of Figure 1 displays two examples of a texture in which a figure can be perceived. The visual system perceives the figure because the line elements in this region have an orientation that is orthogonal to that of the line elements that make up the background. Our visual system groups all the line segments of one orientation into a coherent surface that segregates 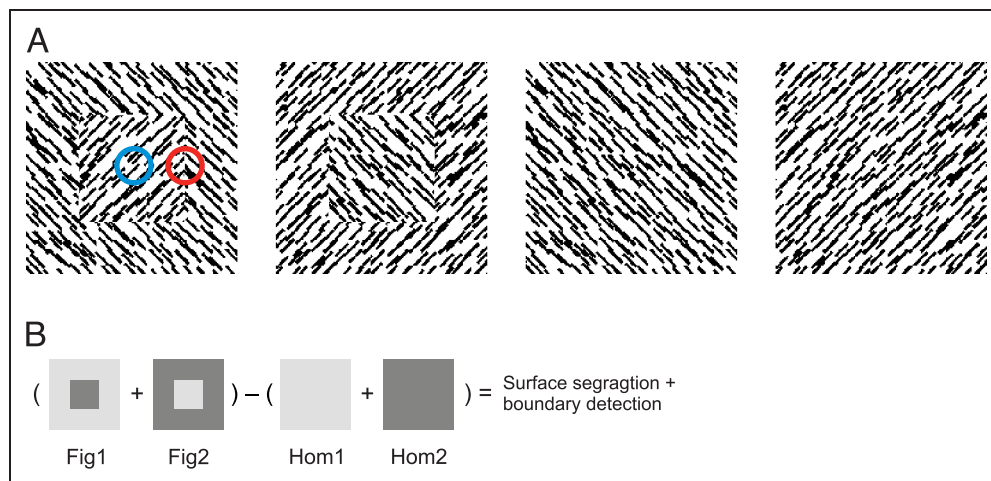from the other line elements (Nothdurft, 1985). A local ori-entation discontinuity encloses the figure surface, forming a boundary, and separates it from the background while the background seems to continue behind the figure. This implies that perceptually the boundary belongs to, and is an intrinsic part of, the figure and not the background (Nakayama, Shimojo, & Silverman, 1989).

Stimuli, like the ones presented in Figure 1, have been used in a large number of studies to investigate aspects of boundary detection and surface segmentation in monkey (Marcus & Van Essen, 2002; Rossi, Desimone, & Ungerleider, 2001; Zipser, Lamme, & Schiller, 1996; Lamme, 1995) and human (Appelbaum, Wade, Vildavski, Petter, & Norcia, 2006; Schira, Fahle, Donner, Kraft, & Brandt, 2004; Altmann, Bulthoff, & Kourtzi, 2003; Kastner, De Weerd, & Ungerleider, 2000; Caputo & Casco, 1999; Reppas, Niyogi, Dale, Sereno, & Tootell, 1997; Lamme, Vandijk, & Spekreijse, 1993; Bach & Meigen, 1992) subjects. In general, such studies use an analysis scheme

---

[1]University of Amsterdam, [2]Netherlands Royal Academy of Sciences
*Present address: University of Exeter, School of Psychology

Figure 1. Studying texture segregation. (A) Typical stimuli used in experiments investigating texture segregation. Stimuli are composed of homogenously oriented line elements. Typically orientations of 45° and 135° are used. Figure stimuli have a region of line elements with an orthogonal orientation compared to the rest of the line elements (e.g., the background). (B) Subtraction scheme for data measured in response to these stimuli to isolate signals related to boundary detection and scene segmentation (in humans, but see Schira et al., 2004) or, depending on the location in the visual field, responses to boundary detection or scene segmentation (in monkey). In monkey it is possible to differentiate between boundary detection and scene segmentation by placing an electrode in V1 in such a way that its receptive field falls on either the texture boundary (red circle) or on the figure surface (blue circle).

in which activity evoked by homogenous textures (two rightmost panels) is subtracted from activity evoked by the texture-defined figures (two leftmost). When focusing on activity in the early visual cortex in monkey, one can thus draw inferences about the timing of boundary detection and scene segmentation because the receptive fields in these areas are small enough to fall entirely in a surface region (blue circle in Figure 1) or fall on a boundary (red circle in Figure 1). Many such studies have found that neurons in early areas such as V1 show stronger responses to boundary segments than to homogenous textures (Marcus & Van Essen, 2002; Rossi et al., 2001; Zipser et al., 1996; Lamme, 1995), or may even respond selectively to the orientation of such boundaries (Sillito, Grieve, Jones, Cudeiro, & Davis, 1995; Grosof, Shapley, & Hawken, 1993). This indicates that these early visual areas are involved in texture boundary detection. Whether these early visual areas also signal the segmentation of the scene into different surfaces is more controversial. Lamme (1995) (see also Zipser et al., 1996) found that V1 neurons respond more strongly to texture elements belonging to a figure surface than to similar elements belonging to a homogenous or background texture, even when the texture boundary is far outside the receptive field (e.g., at the location of the blue circle in Figure 1). This was observed for figure–ground segregation evoked by orientation differences and for motion or stereoscopic depth-defined textures, as well as in various figure–ground organizations (Zipser et al., 1996; Lamme, 1995). They concluded that V1 is involved in surface segregation. Others, however, have challenged this view (Rossi et al., 2001).

The same, or a similar, paradigm has been used in various electroencephalography (EEG) as well as blood oxygenation level-dependent magnetic resonance imaging (BOLD-MRI) studies. MRI studies have indicated the involvement of areas V1, V2, V3, V4, and V3a (Schira et al., 2004; Altmann et al., 2003; Kastner et al., 2000; Reppas et al., 1997) in the differentiation between homogenous texture and figure textures. EEG studies indicate that occipital areas differentiate between homogenous texture and figure textures around 140–219 msec (Caputo & Casco, 1999; Bach & Meigen, 1992; Lamme, Vandijk, & Spekreijse, 1992). These studies found a greater negativity in occipital electrodes for figure textures compared to homogenous textures. However, in these studies, it was not possible to distinguish between signals coming from boundary detection mechanisms and signals related to surface segregation or scene segmentation. The spatial resolution of both EEG and MRI is not high enough to selectively monitor the activity of cells responding to either boundary or figure surface.

Therefore, it is still unknown whether early visual areas such as V1 are critically involved in surface segregation and scene segmentation, or merely contribute to this process by providing signals about texture boundaries to higher areas. To overcome these limitations, we developed stimuli where all texture boundaries were held constant, while manipulating the surface percept from that of a ''frame'' overlying a background to a ''stack'' of surfaces (see Figure 2). Because the texture boundaries are identical for the stack and frame stimuli, bottom–up saliency-based mechanisms cannot detect the differences between the stack and the frame stimuli because this saliency can only be based on the boundary discontinuities (Li, 2001; Sillito et al., 1995).

By comparing neural signals to these ''frame'' and ''stack'' stimuli obtained with EEG and fMRI, and by contrasting these signals with homogeneous textures, it was possible to differentiate between boundary detection and surface segregation signals in humans and to investigate
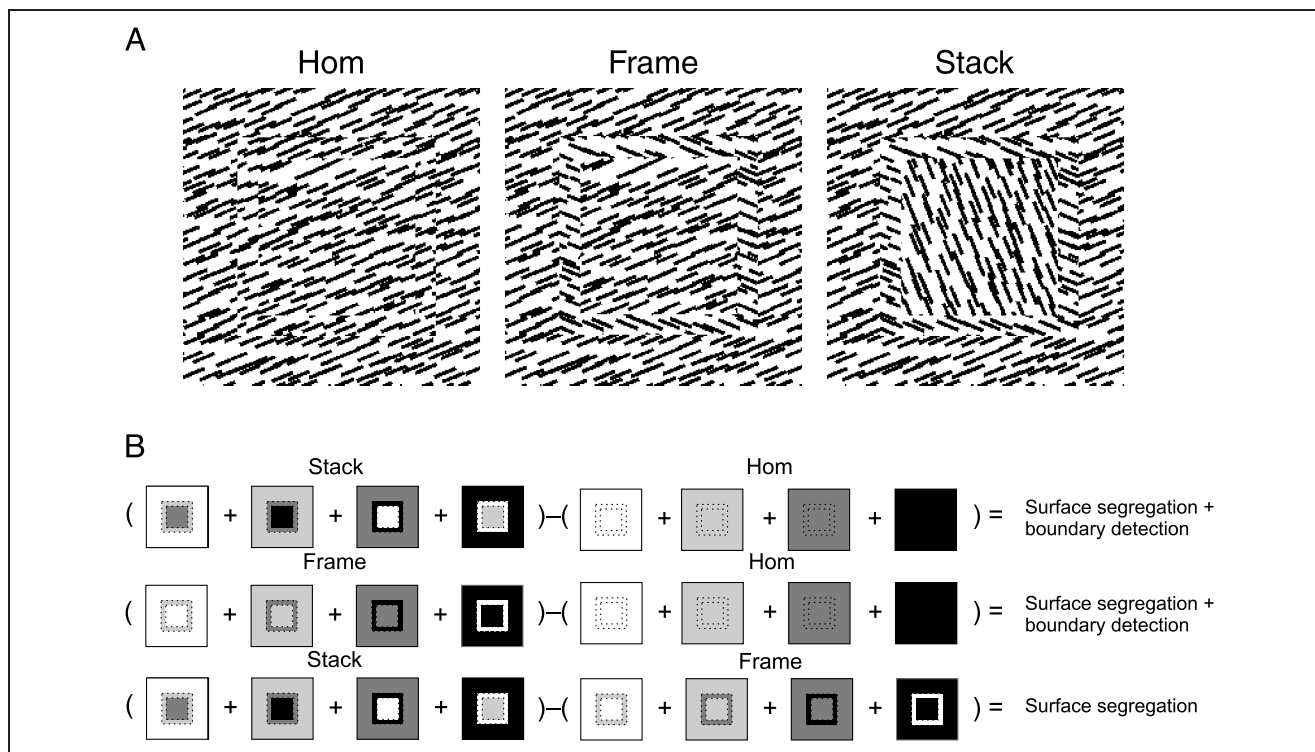
Figure 2. Example of present analysis scheme and stimuli. (A) We presented subjects with stimuli that consisted of homogenously oriented line elements ("Homogenous," left-hand side), homogenously oriented line elements with a superimposed frame ("Frame," middle), and homogenously oriented line elements with a superimposed frame and line elements with a different orientation within the frame ("Stack," right-hand side). (B) Schematic display of the stimuli that were used in the present experiment. Stimuli were generated with different orientations (indicated in gray, either 22.5°, 67.5°, 112.5°, 157.5°) that, when averaged like shown here, were balanced at a local level for local orientation. Furthermore, when responses measured on frame stimuli are subtracted from responses to the stack stimuli, the presence of orientation boundaries is also balanced so that the resulting signals can be attributed to surface segregation mechanisms. Finally, when neural responses do not differ between stack and frame, but do deflect from the homogenous textures, we can relate these signals to pure boundary detection.

whether V1 and other early visual areas are involved in scene segmentation.

METHODS

Twenty-five subjects gave their written informed consent to participate in this study, which was approved by the local ethical committee. Thirteen subjects (13 women, age = 18–24 years) participated in the fMRI experiments and 12 subjects (10 women, age = 18–26 years) participated in the EEG experiments. One subject was removed from the fMRI experiment because it was impossible to identify multiple cortical areas in this subject.

Visual Stimulation: Paradigm

Stimuli consisted of textures with homogenously oriented line elements or textures that contained line elements that formed a frame. The line elements of the frame were always at 45° with those of the background and at 45° with those of the square region enclosed by the frame. The orientation of the background could thus be either identical to that of the enclosed region

("frame") or at 90° with it ("stack") (see top row of Figure 2 and Supplementary Figures 1 to 3). We used four different orientations (22.5°, 67.5°, 112.5°, and 157.5°) to balance these stimuli at a local level of orientation while still being able to render each of the orientations in an identical way on the monitor. The homogenous textures were created by generating two homogenous textures and copying, at the locations of the frames, the content of one of the homogenous textures into the second. These borders are not visible at the size of stimulus presentation.

After counterbalancing using all these orientations, the "frame" and "stack" stimuli have an identical amount of edge, always defined by a 45° orientation difference and, on average, are all defined by the same orientations of line elements.

However, these stimuli differ in terms of the type of segmentation that occurs with the inner area of the frame grouping with the background for the "frame" stimuli while this area segregates for the "stack" stimuli. The internal region of the stacks and frames was 3.6° in width, the surrounding frames was 0.4° wide. A full stimulus screen contained seven "stack" or "frame" regions, with a stimulus in the center, flanked by a

stimulus on the left and the right and, above and below this row, two stimuli centered on the spaces between the stimuli of the middle row. We used seven stimuli to boost the measurable signal. Furthermore, we also presented homogenous (''hom'') textures of 22.5°, 67.5°, 112.5°, and 157.5° with respect to horizontal, containing neither stacks nor frames, but only homogenous texture.

During visual stimulation, we also presented a fixation cross (0.5°) in the center of the screen, and subjects were instructed to fixate on it. The orientation of the fixation cross rotated 3° every 4 sec (on average) and subjects had to indicate when this happened.

Visual Stimulation and Data Analysis: EEG

Stimuli were presented on a 19-in. Ilyama monitor with a resolution of 1024 × 768 pixels and a frame rate of 100 Hz. Subjects were seated 90 cm from the monitor. Stimuli were presented for a period of 500 msec, followed by an interstimulus interval of 500 to 800 msec. Stack, frame, and homogenous textures were presented in randomized order. A session lasted 10 min and two sessions were obtained from each subject. Recordings were made with a Biosemi 48-channel Active Two EEG system (Biosemi Instrumentation BV, Amsterdam, The Netherlands). Data were sampled at 256 Hz. The raw data of a 10-min session were detrended and filtered with a high-pass filter at 0.25 Hz (12 db/octave), a low-pass filter at 30 Hz (48 db/oct), and a notch filter at 50 Hz. After this, we segmented the trials into epochs starting at 250 msec before stimulus onset and ending at 750 msec after stimulus onset. Subsequently, trials with a deflection larger than 250 μV were automatically removed, ocular corrected in which the influence of ocular generated EEG was removed with a regression analysis based on two horizontal and two vertical electrooculogram channels (Gratton, Coles, & Donchin, 1983), and artefact rejected. After this we performed a baseline correction based on the data between −100 and 0 msec relative to stimulus onset, and averaged the trials per condition. To further localize neocortical activity, we performed a spline Laplacian (Perrin, Pernier, Bertrand, & Echallier, 1989) on the resulting average. The spline Laplacian filters against deep cortical sources (Nunez & Srinivasan, 2006). To reduce the multiple comparison problem, we pooled channels that were spatially near to each other into five new leads (see Figure 3, left side) and these where the only leads analyzed.

Data were statistically tested with a paired t test that was performed over the averages of the subjects for the conditions ''stack'' versus ''frame,'' ''stack'' versus ''hom,'' and ''frame'' versus ''hom'' for all samples between 0 and 400 msec. These data were thresholded by means of the false discovery rate (Benjamini & Hochberg, 1995) at a $p < .05$ level. The false discovery rate controls the expected proportion of false positives in a dataset and not the chance of any false positives (as a Bonferroni correction would do). We subsequently looked for the first moment in time at which each of these three pairs of comparisons significantly deflected from zero. This moment was taken as the moment of discrimination.

Visual Stimulation and Data Analysis: fMRI

Stimuli were projected on a screen at the front end of the scanner table (Philips 3T Intera). The projected image was seen via a mirror placed above the subject's head. A magnet-compatible response box was used to record when subjects saw the fixation cross rotate. The subject's head was immobilized using foam pads to reduce motion artifact and earplugs were used to moderate scanner noise.

The different stimulus categories (''stacks,'' ''frames,'' and ''hom'') were presented in blocks that lasted 24 sec, in which stimuli where presented every 2 sec, followed by the presentation of an isoluminant gray stimulus that lasted 16 sec. Each stimulus was presented for 1000 msec and was followed by a gray isoluminant screen. The different categories were presented in pseudorandom order for 12 times per category over three runs (one run lasted for 8 min). BOLD-MRI (GE-EPI, transversal slice orientation, TR = 2295 msec, TE = 28 msec, FOV = 240 mm, matrix size of 96 × 96, slice thickness = 3, slice gap = 0.3, 35 slices, and a sense factor of 2.5) was recorded during the presentation of these blocks.

Data were analyzed with Brainvoyager (Brain Innovation BV, Maastricht, The Netherlands) and Matlab (The Mathworks, Natick, MA). The functional images were motion corrected, slice time aligned, temporally smoothed with a Gaussian filter (FWHH of 2.8 sec), and high-pass filtered (0.01 Hz) in the temporal domain. No spatial smoothing was used. After this, the functional images were aligned to the structural image acquired at the start of each scanning session (T1 turbo field echo, 182 coronal slices, FA = 8, TE = 4.6, TR = 9.7, slice thickness = 1.2, FOV = 256 × 256, matrix = 256 × 256) and transformed, on the basis of this structural image, to Talairach space.

Functional data were normalized toward percent signal change and analyzed by means of a deconvolution analysis (Burock, Buckner, Woldorff, Rosen, & Dale, 1998). Data from areas V1, V2, V3, V3a, V4, lateral occipital (LO), fusiform face area (FFA), parahippocampus place area (PPA), and the middle temporal motion area (MT) (see below for cortical mapping) were pooled into one time course per subject per condition. Activity from these time courses was pooled between 6.9 and 25.2 sec (see Figure 4A). Resulting values were t tested against zero. The percentage signal change was calculated, per subject, for stack versus hom, frame versus hom, and stack versus frame, and we performed a testing over the resulting values against zero.
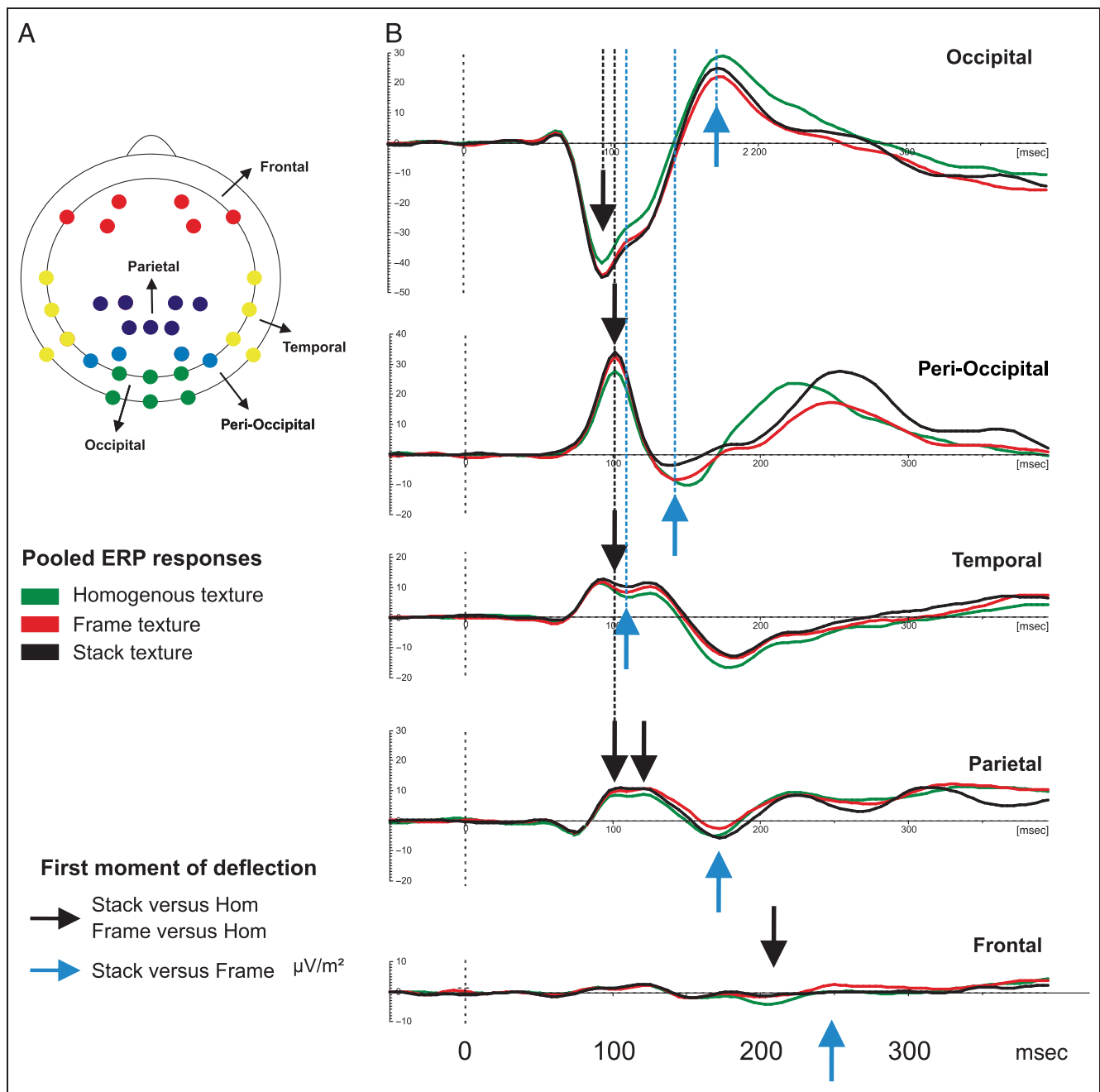
Figure 3. EEG responses. (A) A limited number of ERP channels were pooled into ''local'' leads after applying a spline Laplacian. Colors indicate which electrodes were pooled into which leads. (B) Responses for each of these leads to the homogenous, frame, and stack textures. Bars indicate the first moment of deflection for each of the comparisons (corrected for multiple comparisons with the false discovery rate).

Eye Movements

To force subjects to fixate at the center of the screen, subjects had to report, during the experiment, whenever a fixation cross (0.5°) rotated 3° (this occurred, on average, once every 4 sec). To further control for eye movement, we recorded the eye movements (Resonance Technology/Arrington Research, Scottsdale, AZ) of five subjects during the recording of BOLD-MRI. Eye movements were recorded at a speed of 60 Hz by digitizing video images. Data were analyzed by smoothing the recorded data and classifying visual activity as blinking (i.e., loss of a visible pupil), making a saccade (i.e., an eye movement faster than 30 deg/sec) or fixating.

Cortical Mapping

We presented each of the subjects with stimuli to map the polar angle and eccentricity of their visual BOLD responses in order to determine the locations of areas V1, V2, V3, V3a, and V4. The polar mapping paradigm

Figure 4. (A) BOLD-MRI responses. Grand-average response in area V1 for each of the three stimuli. Error bars indicate standard errors. The area under the curve within the gray area was used to calculate, per subject, the response toward the stimulus. (B) Results from the area under the curve pooling for each of the cortical areas under consideration. Error bars indicate *SE*.
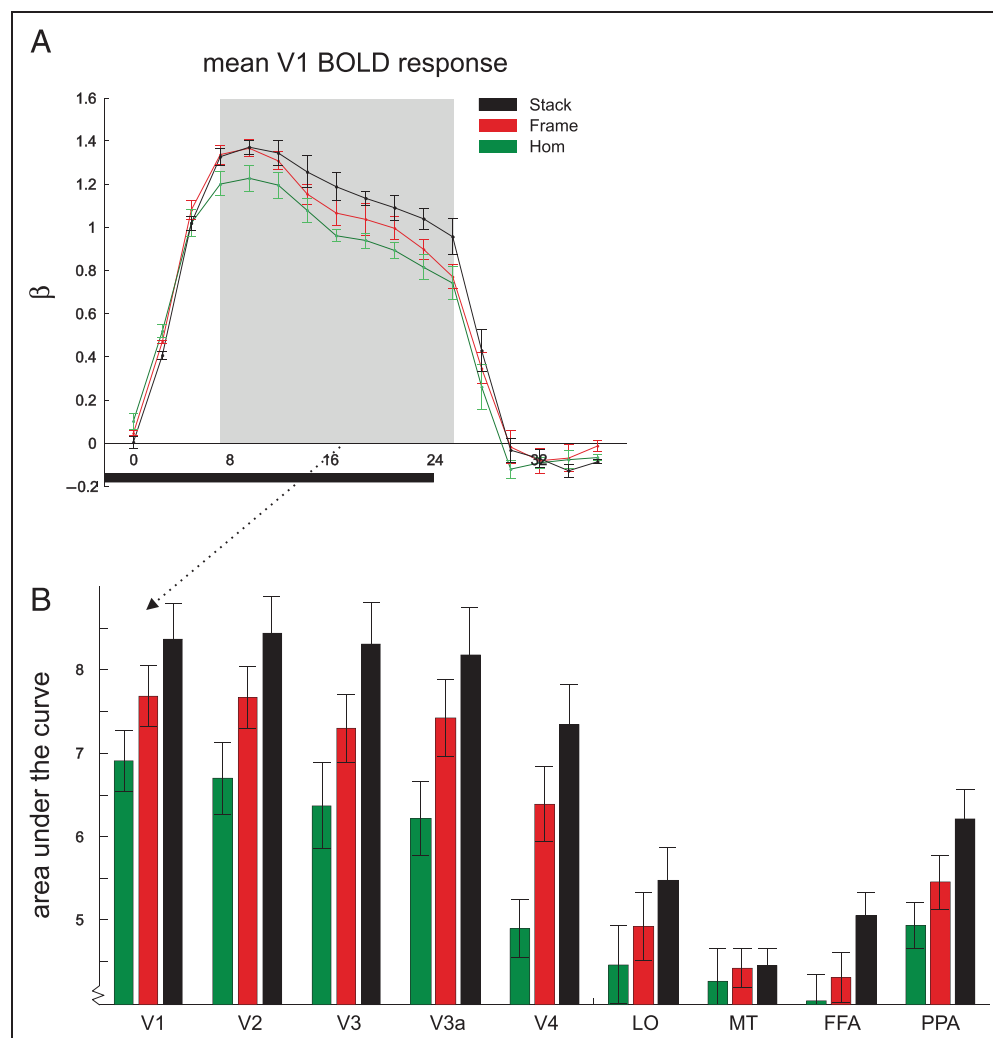
consisted of a rotating wedge made of red and green squares (wedge of 30° that rotated at a speed of 6.4 deg/sec and extended 8° in space, with a 6-Hz contrast polarity reversal) that rotated clockwise or counterclockwise for a total period of 494 sec.

The eccentricity mapper consisted of an expanding annulus made of red and green squares with a width of 1.2° at 4°. The exact size of the annulus at other eccentricities was determined on the basis of the cortical magnification factor (Dougherty et al., 2003), and the annulus dilated or expanded at a speed of 0.33 deg/sec, with a 6-Hz contrast polarity reversal for a total period of 320 sec.

BOLD-MRI was recorded during the presentation of each of these paradigms (GE-EPI, transversal slice orientation, TR = 2000 msec, TE = 28 msec, FOV = 200 mm, matrix size of 112 × 112, slice thickness 2.5, slice gap 0.3, 24 slices, and a sense factor of 2.5). The data from these mappers were analyzed by correlating the BOLD responses with a model of the hemodynamic response function (HRF) at different lags (Linden, Kallenbach, Heinecke, Singer, & Goebel, 1999) and projecting the lag information on a reconstruction of the cortical sheet.

This reconstruction was made with Brainvoyager on the basis of two recordings of a high-resolution MRI scan (T1 turbo field echo, 182 coronal slices, FA = 8, TE = 4.6, TR = 9.7, slice thickness = 1.2, FOV = 256 × 256, matrix = 256 × 256). The borders between V1, V2, V3, V4, and V3a were subsequently manually drawn in (Sereno et al., 1995).

Area MT was mapped by presenting blocks of coherently moving dots and comparing activity evoked in these block with presentations of randomly appearing dots while measuring BOLD-MRI (same as above). Each of these blocks lasted for a period of 16 sec and the total presentation lasted for 320 sec. During a 16-sec block of coherent movement, dots alternated, moving inward and outward every 2 sec. Data were analyzed by means of a general linear model (GLM). Predictors were generated by convolving the onset times of the moving stimuli and nonmoving stimuli with a model of the HRF model, fitting these to the MRI data (AR(1) autocorrelation correction), generating a contrast between these two predictors and projecting the resulting data on the surface-based reconstruction of that subject (thresholded at a $t = 4$ level). We determined area MT to be

the activation cluster in the inferior temporal sulcus and the lateral occipital sulcus (Dumoulin et al., 2000).

Areas LO, FFA, and PPA were mapped by presenting blocked presentations of faces, houses, novel objects (bottles, chairs, and scissors) of different orientations and phase scrambled versions of these objects while measuring BOLD-MRI (same as above, except 28 slices). Subjects performed a 1-back task. Stimuli were presented every 2 sec in blocked presentations that lasted 16 sec followed by a baseline of 16 sec and each block was repeated four times. Predictors were made by convolving the onset times of the stimuli from the different categories with a model of the HRF and fitting these to the data with the GLM. To determine the location of FFA, we contrasted faces versus houses and novel objects; to determine the location of PPA, we contrasted houses versus faces and novel objects; and to determine the location of LO, we contrasted houses, faces, and novel objects versus the scrambled versions of these objects (Grill-Spector & Malach, 2004). Area PPA was considered to be the activation in the parahippocampal gyrus, area FFA was considered to be the activation in the fusiform gyrus, and area LO was considered to be the activation between area MT and the early retinotopic areas. We failed to localize area PPA in one subject and area LO in another subject.

RESULTS

Behavior

The reaction time (RT) of the subject on the primary fixation task (detecting a rotation in the fixation cross) was identical for the situation in which homogenous (mean = 547 msec, SD = 83 msec), frame (mean = 554 msec, SD = 85 msec), and stack textures (mean = 550 msec, SD = 77) were presented. Detection performance was 99% (SD = 0.01%) in the situation in which a homogenous texture was presented and 97% (SD = 0.01%) in cases where a stack or frame texture was presented. Detection of the rotation in the fixation cross was significantly less in the situation where stack (t = 3.7, df = 23, p = .001) or frame textures (t = 4.15, df = 23, p = .0004) were presented compared to the presentation of homogenous textures. This means that the presence of the stack and frame patterns interferes with the detection task. Given that RT on the rotation detection time did not differ between homogenous textures, and the performance difference is only 2%, we interpret this to mean that the presence of the frame and homogenous textures interfered with the rotation detection task in a small minority of trials.

Electroencephalography

The stack texture was comparable to the frame texture in the sense that two parallel texture boundaries were present. However, because the orientation of the region enclosed by the frame was orthogonal to that of the background, the percept evoked by this stimulus was that of a "stack" of texture squares. A key feature of the current stimuli is that both the stack and the frame texture have texture boundaries at the exact same locations.

Moreover, in both cases, these two boundaries are formed by a 45° difference in orientation. The only difference between the two is at the level of the global surface percept (frame vs. stack). The enhanced saliency of the "stack" stimuli compared to the "frame" stimuli is therefore, in all likelihood, a (indirect) result of the difference in segmentation.

The responses evoked by the hom, frame, and stack textures were transformed with a spline Laplacian and pooled into five predetermined clusters. We determined the first moment at which each of these conditions deflected from the other conditions by means of a paired t test over subjects and thresholding the result with FDR at a p < .05 level. Results indicate that responses on the stack and frame textures start to deflect from the responses evoked by the homogenous textures at 92 msec after stimulus onset in the occipital pooling.

This same pattern can be observed at 104 msec in the peri-occipital and temporal pooling. The parietal pooling shows a more complicated pattern because the responses evoked by the stack textures start to deflect at 104 msec from the responses evoked by the hom textures, whereas this only occurs at 120 msec for the pair frame versus homogenous. Taken that all other poolings show an identical moment of deflection between stack and frame on the one hand and homogenous on the other, we believe that it is likely that boundary detection takes place in this area at 104 msec and that the later deflection between frame and homogenous is caused by noise.

Finally, at 208 msec, a deflection at the frontal electrodes between the responses evoked by the stack and frame stimuli versus the responses evoked by the homogenous textures (see Figure 3B and Table 1) can be detected.

When the frame and stack textures start to differentiate from the homogenous textures, the responses are consistent with earlier studies (Appelbaum et al., 2006; Caputo & Casco, 1999; Bach & Meigen, 1992; Lamme et al., 1992) and are always larger. A more complex pattern emerges when stack and frame textures start to differentiate from each other. When this occurs in the occipital, temporal, and parietal poolings, the stack textures evoke a stronger response than the frame texture. For the peri-occipital and frontal channels, these responses show the opposite pattern. The most attractive explanation for larger responses in the event-related potential (ERP) would be the presumption of a higher degree of underlying neural activity. However, because we measure activity resulting from the addition of multiple underlying neural sources with potential different

Table 1. Moment of First Deflection between Different Conditions for the ERP Responses

Pooling	Comparison	Moment of First Deflection (msec)
Occipital	Stack/Hom	92
	Frame/Hom	92
	Stack/Frame	172
Peri-Occipital	Stack/Hom	104
	Frame/Hom	104
	Stack/Frame	140
Parietal	Stack/Hom	104
	Frame/Hom	120
	Stack/Frame	172
Frontal	Stack/Hom	208
	Frame/Hom	208
	Stack/Frame	248
Temporal	Stack/Hom	104
	Frame/Hom	108
	Stack/Frame	112

This moment was determined by means of a paired t test and the results of this test where thresholded by means of a false discovery rate. Finally, we searched for each pair of comparisons the first moment at which this comparison deflected from zero.

polarities, this can never be a strong inference (Luck, 2005). We can conclude that the brain differentiates, at these moments in time, between stack and frame stimuli.

The difference between responses evoked by stack textures versus responses evoked by frame textures shows a remarkably different pattern of activation. The stack and frame textures have the same amount of edge but have a different scene organization. A difference

between these two conditions can first be observed in the temporal pooling at 112 msec, followed by the peri-occipital pooling at 140 msec and the parietal and occipital pooling at 172 msec. Finally, the frontal pooling shows a deflection between these two conditions after 248 msec (see Figure 3B and Table 1).

Functional Magnetic Resonance Imaging

The same stimuli used in the EEG experiment were also used in the fMRI experiment. If early visual processing is limited to boundary detection, there should be no difference in the responses to stack and frame textures. If, on the other hand, early visual processing is also related to global scene segmentation, it is expected that we find response differences in early visual areas.

The area under the curve of the response toward each of the stimuli was calculated for each cortical area. The resulting values are shown in Figure 4B. We also calculated, per cortical area, the percentage signal change for the pairing stack versus homogenous, frame versus homogenous, and stack versus frame (see Figure 5, Table 2). A paired t test was used for the statistics presented in Table 2 and Figure 5.

Results indicate that areas V1 to V4 and V3a differentiate between all the different stimulus categories and therefore show that these areas are not only involved in boundary detection but also reflect to the process of scene segmentation. The difference between stack and frame textures on the one hand and homogenous texture on the other (blue and red bars in Figure 5) increases from V1 to higher-tier areas and is significant for the comparisons between area V1 versus V4 for stack versus homogenous ($t = 6.52, df = 11, p = .000$) and frame versus homogenous ($t = 3.3, df = 11, p = .007$). These response differences are mainly caused by a decrease in the response toward the homogenous texture (see Figure 4, green bars).

Figure 5. Perceptual difference between responses (signal change) of stack versus hom textures, frame versus hom textures, and stack versus frame textures. Error bars indicate the variance of these responses based on the percentage signal change per subject.
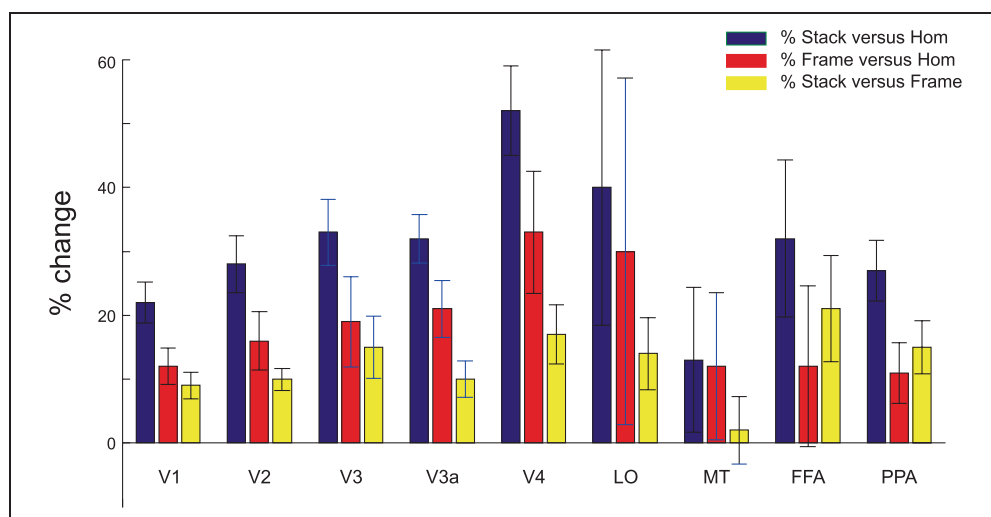
Table 2. Results from BOLD-MRI Comparisons

Area	Comparison	t	df	Significance (two-tailed)
V1	Stack/Hom	7.02	11.00	.00002*
	Frame/Hom	4.39	11.00	.00108*
	Stack/Frame	4.41	11.00	.00104*
V2	Stack/Hom	6.50	11.00	.00004*
	Frame/Hom	3.73	11.00	.00335*
	Stack/Frame	6.08	11.00	.00008*
V3	Stack/Hom	6.73	11.00	.00003*
	Frame/Hom	2.79	11.00	.01770*
	Stack/Frame	3.24	11.00	.00787*
V4	Stack/Hom	8.77	11.00	.00000*
	Frame/Hom	4.92	11.00	.00046*
	Stack/Frame	3.85	11.00	.00270*
V3a	Stack/Hom	7.74	11.00	.00001*
	Frame/Hom	3.64	11.00	.00389*
	Stack/Frame	3.67	11.00	.00368*
LO	Stack/Hom	1.85	10.00	.09360
	Frame/Hom	1.12	10.00	.29053
	Stack/Frame	2.43	10.00	.03528*
MT	Stack/Hom	1.22	11.00	.24795
	Frame/Hom	1.13	11.00	.28376
	Stack/Frame	0.39	11.00	.70197
FFA	Stack/Hom	2.73	11.00	.01960*
	Frame/Hom	1.04	11.00	.32124
	Stack/Frame	2.68	11.00	.02140*
PPA	Stack/Hom	5.66	10.00	.00021*
	Frame/Hom	2.37	10.00	.03907
	Stack/Frame	3.61	10.00	.00479*

Comparisons where performed with a paired *t* test for all pairs of stimuli per cortical area.

*FDR(0.05) = 0.0353.

Activity related to the stack–frame contrast shows a very different pattern of activity. The percent signal change between the stack and frame textures (yellow bars, Figure 5) is much more constant and shows no significant differences between early visual areas.

Interestingly, we find that areas PPA and FFA both differentiate between stack versus homogenous textures and frame versus stack textures but fail to show this difference for the comparison frame versus hom (see Table 2), indicating that these areas respond reliably to aspects of the scene organization but do not respond to the presence of edges.

Area LO shows, in terms of response magnitude, a pattern that is similar to area V4, albeit with a higher within-subject variance on the homogenous textures. This area does respond significantly different between presentations of stack textures and presentations of frame textures (see Table 1).

In sharp contrast, and included in this experiment as a negative control, area MT showed no significant difference between any of the stimulus categories.

Eye Movement and Attention Controls

One possible confound, when comparing BOLD-MRI activity measured in visual areas, is that a difference in response is obtained because subjects make more saccades, or have a different pattern of eye movements or fixation, in one of the conditions, or that attention is selectively drawn to one of the conditions. To control for this, all subjects performed a task in which they had to detect small changes in the orientation of a cross at the center of the screen, which requires accurate fixation and attention. Overall performance for subjects on this task was 96% (SD = 2%). Furthermore, we recorded the eye movements of five subjects during the main experiment and analyzed the number of saccades each of the subjects made. Results show that these did not differ significantly between the different conditions: 6 saccades (SD = 3.8) during presentations of homogenous textures, 7.8 saccades (SD = 2.8) during presentations of frame textures, and 5.3 saccades (SD = 2.4) during presentations of stack textures.

DISCUSSION

Boundary Detection versus Surface Segregation

The stack texture was comparable to the frame texture in the sense that two parallel texture boundaries were present. However, because the orientation of the region enclosed by the frame was orthogonal to that of the background, the percept evoked by this stimulus was that of a "stack" of texture squares, with the center one somewhat smaller than the one below. The stimuli used in the current experiment therefore made it possible to differentiate between the processes of texture boundary detection and surface segregation or filling-in: a difference in response between stack or frame and the homogenous stimuli is caused by both boundary detection and surface segregation mechanisms, but a difference between stack and frame is only caused by surface segregation mechanisms. It seems logically warranted that boundary detection precedes surface segregation. However, this is also supported by earlier findings in monkey V1, showing boundary detection at 90 msec and surface segregation at 120 msec (Lamme, Rodriguez-Rodriguez, & Spekreijse, 1999). Therefore, we can follow the temporal evolution of both processes from our EEG results,

assuming that the earliest difference between stack–frame and homogenous will be the result of boundary detection, whereas the earliest difference between stack and frame will be the result of surface segregation.

Occipital channels reflect boundary detection at 92 msec after stimulus onset, followed by peri-occipital (104 msec), temporal (104–108 msec), and parietal (104–120 msec) channels (Figure 3 and Table 1). This is fully consistent with a feedforward spread of information, and boundaries between textures being detected at successive levels of the visual hierarchy (Roelfsema, Lamme, Spekreijse, & Bosch, 2002; Lamme & Roelfsema, 2000). Our fMRI results additionally show that boundary signals get stronger as they proceed through the visual hierarchy (Figure 5, blue bars, red bars).

Surface segregation, on the other hand, shows a radically different temporal evolution. The earliest stack versus frame difference is found in the temporal channels at 112 msec, that is, very briefly after the boundary detection signals have reached this area. From there, surface segregation signals spread toward the peri-occipital (140 msec), parietal (172 msec), and occipital (172 msec) channels (Figure 3, Table 1). Note that the electrodes receiving signals from the areas that are lowest in the visual hierarchy (occipital channels) reflect surface segregation at the longest latency. This temporal ordering is consistent with surface segregation depending on a feedback process, where activity first arises in high-level visual areas, and then spreads toward lower visual areas. Another difference between boundary detection and surface segregation signals is that boundary detection signals increase in strength as one progresses through the visual hierarchy, whereas surface segregation signals are about equally strong in all visual areas, as measured with BOLD-fMRI (Figure 5, yellow bars).

Both findings, a feedforward detection of texture boundaries, followed by surface segregation mediated by feedback connections, are fully consistent with a neural network model of scene segmentation that employs feedforward, horizontal, and feedback connections in the visual cortex (Roelfsema et al., 2002). This model replicated earlier neurophysiological findings of boundary detection and surface segregation as found in V1 of the monkey (Lamme et al., 1999; Lamme, Super, & Spekreijse, 1998). Our current findings further corroborate this model by replicating the temporal order of events in other areas as well.

Our findings are roughly consistent with a study by Caputo and Casco (1999), who parametrically varied the visibility (figure elements could be orthogonal or parallel with figure direction, which affect perceptual saliency without affecting boundary layout) of the presented figure to disentangle boundary detection and surface segregation. They showed that ERP components around 140–160 msec in occipital areas are more related to the processing of boundary detection and components around 200–260 msec are more related to scene segmen-

tations. However, the timings of the critical events in the current study are generally earlier (92 and 172 msec) than in their study.

Another study from our lab, using the same stack–frame stimulus design, showed that late feedback signals are, indeed, necessary for a normal figure–ground percept. In that study, transcranial magnetic stimulation (TMS) was applied to the occipital lobe at various delays after stimulus presentation (Heinen, Jolij, & Lamme, 2005). It was shown that TMS pulses targeted at V1, yet applied as late as 250 msec after stimulus onset, could still disrupt the subjects' ability to discriminate between stack and frame organizations.

Some further points are worth noting. First, our data confirm the validity of the stack–frame manipulation. A potential concern is that the frame width is not large enough to prevent neurons in, for example, area V1 to ''see'' both the area surrounding the frame and the region within the frame in one receptive field. In that case, a stack–frame difference, which we here attribute to surface segregation mechanisms, might simply be caused by local orientation discontinuity detectors (e.g., Sillito et al., 1995). However, the finding of a considerably longer latency in the occipital channels for stack versus frame compared to stack (or frame) versus homogenous argues against that. In addition, such an explanation would have predicted the strongest stack versus frame signals in higher visual areas, where receptive fields are largest. That is also not what we found (Figure 5, yellow bars).

Second, our BOLD-MRI recordings indicate that areas V1, V2, V3, V3a, and V4 all respond differentially toward the three different types of stimuli. This indicates that all early visual areas are involved in both boundary detection and scene segmentation. The responses that we measured in higher visual areas show a more complicated pattern. We find a consistent ordering of the weakest response on the homogenous textures, a stronger response on the frame textures, and the strongest response on the stack textures in areas LO, FFA, and PPA. All of these areas differentiated between frame and stack textures, indicating that they signal differences in scene organization. Areas FFA and PPA differ in their response between stack and homogenous textures, but not in their response to frame and homogenous textures. This somewhat curious response behavior is a result of a much smaller within-subject variance for the responses between the stack texture and the frame textures compared to the homogenous texture. No consistent response was found in area MT (Table 2 and Figure 5). Taken together, these data are consistent with our EEG findings, although with BOLD-MRI we cannot disentangle the temporal order of activation of areas. For example, the MRI data support our conclusion that neural activity distinguishing between stack and frame stimuli, and thus, related to scene segmentation, indeed is present in early visual areas such as V1. Our EEG

results add to this that this activity arises later in time than similar activity found in higher areas in the ventral pathway, thus supporting the idea that scene segmentation depends on feedback signals.

Third, at a considerably longer latency (208–248 msec), both boundary detection and surface segregation signals are expressed in frontal regions (Figure 3, Table 1). We can only speculate about the function of these signals, but earlier studies suggest that this may reflect attention or other cognitive processes being guided toward the texture-defined objects. The precise origin of these signals is also not clear, as frontal areas were not included in our fMRI recordings.

Surface Segregation Signals in V1: The Controversy

As stated in the Introduction, one of the main controversies in this field is whether V1 neurons signal figure–ground organization and scene segmentation or not. This hypothesis was coined on the basis of earlier findings in monkey. It was found that a V1 neuron responds more strongly when its receptive field is inside a figure region than when it is stimulated by identical background elements, even though its receptive field is much smaller than the figure (Marcus & Van Essen, 2002; Lamme et al., 1998; Zipser et al., 1996; Lamme, 1995). This was found not only for orientation defined figures (as used here) but also when segregation was based on differences in motion, stereo-depth, or color (Zipser et al., 1996; Lamme, 1995). Also, the response modulation was consistent with the perceived figure–ground organization in a variety of scene configurations (Zipser et al., 1996). Therefore, these "contextual modulations" have been interpreted as playing a role in figure–ground segregation (Zipser et al., 1996; Lamme, 1995), or at least as reflecting neural processes related to grouping (Kapadia, Ito, Gilbert, & Westheimer, 1995) and segregation (Nothdurft, Gallant, & Van Essen, 1999; Knierim & van Essen, 1992). A recent fMRI study has lent support to these ideas by showing that V1 (together with other early visual areas) is involved in the grouping of contour elements into coherent shapes (Altmann et al., 2003).

Contextual modulation typically occurs at some delay with respect to the visual response itself. This delay may vary from 10 to about 60 msec, depending on the nature of the surround stimulation (Lamme & Spekreijse, 2000). This suggests that contextual modulation is not mediated by a strictly feedforward process, but involves either horizontal interactions within visual areas, or feedback connections from higher to lower areas (Lamme & Roelfsema, 2000). The importance of horizontal connections is indicated by the similarity between contextual effect and the layout of these connections (Das & Gilbert, 1999). The role of feedback has been indicated by lesion studies, showing the contribution of higher

areas in figure–ground signals in V1 (Bullier, Hupe, James, & Girard, 2001; Lamme et al., 1998). In addition, it has been demonstrated that the conduction speeds of lateral connections are not fast enough to integrate information beyond the classical receptive field and its proximal surround (Angelucci et al., 2002). Furthermore, the cortical magnification factor of the visual cortex provides a "natural" experiment for the involvement of lateral connections. If lateral connections mediate contextual effects at a relative large scale, it would be expected that these effects would be measured later for more centrally presented stimuli (where lateral connections need to travel a large distance) and earlier for more peripheral presented stimuli. Because there is only a modest dependency on cortical distance (Bair, Cavanaugh, & Movshon, 2003), the most consistent interpretation is that contextual modulation at some distance from the classical receptive field is caused by recurrent processing. This leaves feedback processing as the only explanation for correlates of scene segmentation in area V1, especially because these have an adequate conduction speed (3.5 m/sec) comparable to the speed of feedforward connections (Bullier et al., 2001).

Other studies have, however, challenged these ideas. Rossi et al. (2001) did not find any significant contextual modulation of V1 neurons when their receptive field was not in the vicinity of a texture boundary. An early fMRI study (Kastner et al., 2000) did not find significant activation of V1 related to texture segregation, and more recently, Schira et al. (2004) tried to disentangle boundary detection signals from surface segregation signals by selectively looking at MRI voxels that represent texture boundaries or those that represent figure surfaces. They found boundary signals in V1, but no surface segregation activity. So the question remains: is V1 involved in grouping and segmentation or not? The differences between studies might be attributable to two factors: signal-to-noise ratio and attention.

Signal Strength and the Role of Attention

We believe that the apparent difference between the current data and the fMRI studies finding no responses in early visual areas is only quantitative. We find the highest levels of activation for the contrast stack versus homogenous texture in areas V4 and V3a, whereas this difference is much smaller in early visual areas. This is similar to the pattern of activation found in the study of Kastner et al. (2000), the only difference being that, in that study, the V1 and V2 signals were not significant. In general, the studies finding no contribution of area V1 have used a substantial smaller number of subjects than either the present study or the Altmann et al. (2003) study. We therefore believe that our positive findings are not invalidated by these negative findings.

A second issue is attention, or awareness of the figure patch. Both Kastner et al. (2000) and Schira et al. (2004)

took measures to make sure that their subjects were not aware of the different stimuli they used. This also prevented the subjects from focusing their attention on the figure patches. In our study, subjects were aware of the different stimulus configurations (in the sense that they were told in advance about them). However, they were required to perform a demanding foveal task, preventing attention from being drawn by the different figure configurations.

These differences between studies could argue for attention (or rather awareness) being the source of the signals recorded. Surface segregation signals may alternatively be interpreted as reflecting attention drawn by these surfaces. Several findings argue against such an interpretation, however. Figure–ground modulation in monkeys is equally strong when one, two, or four figures are present in the display (Landman, Spekreijse, & Lamme, 2003; Lamme et al., 1998). Modulation does not seem to suffer from attentional bottlenecks or competition. Also, selectively attending to one of two figures does not influence these signals (Marcus & Van Essen, 2002).

Our present findings are mostly ambiguous in this respect because we did not manipulate attention explicitly. However, we did present subjects with a distracting fixation task and subjects were not informed about the relevance or importance of the different textures, nor did they have to respond toward these textures. We therefore do not believe that the various stimuli used in the current study (figures, frames, stacks) differed dramatically in terms of the amount of spatial attention that they drew. If anything, the results may be confounded by differences in object-based attention (because the objects had different shapes but were at the same locations).

Having said that, there is strong evidence from a similar experiment, where we did manipulate attention, that attention is not the primary cause of the signals recorded here. We recently recorded "figure" versus "homogenous" texture EEG and BOLD responses and compared signals in an inattentional blindness condition with signals during normal perception of figures. We had subjects engage in an attentionally demanding task and did not inform them of the presence of texture figures. When tested afterward, we confirmed the full absence of memory or awareness of the figures that were presented in a number of subjects (~50%), and analyzed only the signals from those subjects. We found a significant response difference between figure and homogenous textures in both the inattentional blindness and normal perception condition before 200 msec and early cortical areas. Surprisingly, the size of this response difference before 200 msec, and in early visual areas, was identical for the inattentional blindness and normal perception conditions (Scholte, Witteveen, Spekreijse, & Lamme, 2006), suggesting that most of the segmentation signal (hence, most of the surface segregation

signal) is also present when subjects are not aware and not attending to the figures and is not influenced (before 200 msec) by top–down attention mechanisms. This observation was recently confirmed by Heinrich, Andres, and Bach (2007).

In summary, therefore, we think that the surface segregation signals found here in V1 should not be interpreted as reflecting (object-based) attention but rather as a purely visual processing phenomenon. Awareness, or something like expectancy, might, however, play a role in the presence or strength of these signals. This is at least suggested by another study in monkeys, where perceptual (not reportable) awareness was shown to be a determining factor in the occurrence of contextual modulation in V1 (Super, Spekreijse, & Lamme, 2001). A study using the current paradigms to dissociate boundary detection from surface segregation, while attention is manipulated explicitly, may shed further light on this complex issue (for further discussion on the different roles of attention and awareness, see Lamme, 2004).

REFERENCES

Altmann, C. F., Bulthoff, H. H., & Kourtzi, Z. (2003). Perceptual organization of local elements into global shapes in the human visual cortex. *Current Biology, 13,* 342–349.

Angelucci, A., Levitt, J. B., Walton, E. J., Hupe, J. M., Bullier, J., & Lund, J. S. (2002). Circuits for local and global signal integration in primary visual cortex. *Journal of Neuroscience, 22,* 8633–8646.

Appelbaum, L. G., Wade, A. R., Vildavski, V. Y., Petter, M. W., & Norcia, A. M. (2006). *Journal of Neuroscience, 26,* 11695–11807.

Bach, M., & Meigen, T. (1992). Electrophysiological correlates of texture segregation in the human visual evoked potential. *Vision Research, 32,* 417–424.

Bair, W., Cavanaugh, J. R., & Movshon, J. A. (2003). Time course and time–distance relationships for surround suppression in macaque V1 neurons. *Journal of Neuroscience, 23,* 7690–7701.

Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate—A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society, Series B: Methodological, 57,* 289–300.

Bullier, J., Hupe, J. M., James, A. C., & Girard, P. (2001). The role of feedback connections in shaping the responses of visual cortical neurons. *Progress in Brain Research, 134,* 193–204.

Burock, M. A., Buckner, R. L., Woldorff, M. G., Rosen, B. R., & Dale, A. M. (1998). Randomized event-related experimental designs allow for extremely rapid presentation rates using functional MRI. *NeuroReport, 9,* 3735–3739.

Caputo, G., & Casco, C. (1999). A visual evoked potential correlate of global figure–ground segmentation. *Vision Research, 39,* 1597–1610.

Das, A., & Gilbert, C. D. (1999). Topography of contextual modulations mediated by short-range interactions in primary visual cortex. *Nature, 399,* 655–661.

Dougherty, R. F., Koch, V. M., Brewer, A. A., Fischer, B., Modersitzki, J., & Wandell, B. A. (2003). Visual field representations and locations of visual areas V1/2/3 in human visual cortex. *Journal of Vision, 3,* 586–598.

Dumoulin, S. O., Bittar, R. G., Kabani, N. J., Baker, C. L., Le Goualher, G., Pike, G. B., et al. (2000). A new anatomical landmark for reliable identification of human area V5/MT: A quantitative analysis of sulcal patterning. *Cerebral Cortex, 10,* 454–463.

Gratton, G., Coles, M. G., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology, 55,* 468–484.

Grill-Spector, K., & Malach, R. (2004). The human visual cortex. *Annual Review of Neuroscience, 27,* 649–677.

Grosof, D. H., Shapley, R. M., & Hawken, M. J. (1993). Macaque-V1 neurons can signal illusory contours. *Nature, 365,* 550–552.

Heinen, K., Jolij, J., & Lamme, V. A. F. (2005). Figure–ground segregation requires two distinct periods of activity in VI: A transcranial magnetic stimulation study. *NeuroReport, 16,* 1483–1487.

Heinrich, O., Andres, M., & Bach, M. (2007). Attention and visual texture segregation. *Journal of Vision, 7,* 1–10.

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.

Kastner, S., De Weerd, P., & Ungerleider, L. G. (2000). Texture segregation in the human visual cortex: A functional MRI study. *Journal of Neurophysiology, 83,* 2453–2457.

Knierim, J. J., & van Essen, D. C. (1992). Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. *Journal of Neurophysiology, 67,* 961–980.

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

Lamme, V. A. F. (2004). Separate neural definitions of visual consciousness and visual attention: A case for phenomenal awareness. *Neural Networks, 17,* 861–872.

Lamme, V. A. F., Rodriguez-Rodriguez, V., & Spekreijse, H. (1999). Separate processing dynamics for texture elements, boundaries and surfaces in primary visual cortex of the macaque monkey. *Cerebral Cortex, 9,* 406–413.

Lamme, V. A. F., & Roelfsema, P. R. (2000). The distinct modes of vision offered by feedforward and recurrent processing. *Trends in Neurosciences, 23,* 571–579.

Lamme, V. A. F., & Spekreijse, H. (2000). Contextual modulation in V1 and scene perception. In M. Gazzaniga (Ed.), *The cognitive neurosciences*. Cambridge: MIT Press.

Lamme, V. A. F., Super, H., & Spekreijse, H. (1998). Feedforward, horizontal, and feedback processing in the visual cortex. *Current Opinion in Neurobiology, 8,* 529–535.

Lamme, V. A. F., Vandijk, B. W., & Spekreijse, H. (1992). Texture segregation is processed by primary visual cortex in man and monkey—Evidence from Vep experiments. *Vision Research, 32,* 797–807.

Lamme, V. A. F., Vandijk, B. W., & Spekreijse, H. (1993). Contour from motion processing occurs in primary visual cortex. *Nature, 363,* 541–543.

Landman, R., Spekreijse, H., & Lamme, V. A. F. (2003). Large capacity storage of integrated objects before change blindness. *Vision Research, 43,* 149–164.

Li, B., Chen, Y., Li, B. W., Wang, L. H., & Diao, Y. C. (2001). Pattern and component motion selectivity in cortical area PMLS of the cat. *European Journal of Neuroscience, 14,* 690–700.

Linden, D. E. J., Kallenbach, U., Heinecke, A., Singer, W., & Goebel, R. (1999). The myth of upright vision. A psychophysical and functional imaging study of adaptation to inverting spectacles. *Perception, 28,* 469–481.

Luck, S. J. (2005). *An introduction to event-related potential techniques*. Cambridge: MIT Press.

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.

Nakayama, K., Shimojo, S., & Silverman, G. H. (1989). Stereoscopic depth: Its relation to image segmentation, grouping, and the recognition of occluded objects. *Perception, 18,* 55–68.

Nothdurft, H. C. (1985). Orientation sensitivity and texture segmentation in patterns with different line orientation. *Vision Research, 25,* 551–560.

Nothdurft, H. C., Gallant, J. L., & Van Essen, D. C. (1999). Response modulation by texture surround in primate area V1: Correlates of ''popout'' under anesthesia. *Visual Neuroscience, 16,* 15–34.

Nunez, P., & Srinivasan, R. (2006). *Electric fields of the brain: The neurophysics of EEG* (2nd ed.). Oxford: Oxford University Press.

Perrin, F., Pernier, J., Bertrand, O., & Echallier, J. F. (1989). Spherical splines for scalp potential and current density mapping. *Electroencephalography and Clinical Neurophysiology, 72,* 184–187.

Reppas, J. B., Niyogi, S., Dale, A. M., Sereno, M. I., & Tootell, R. B. (1997). Representation of motion boundaries in retinotopic human visual cortical areas. *Nature, 388,* 175–179.

Roelfsema, P. R., Lamme, V. A. F., Spekreijse, H., & Bosch, H. (2002). Figure–ground segregation in a recurrent network architecture. *Journal of Cognitive Neuroscience, 14,* 525–537.

Rossi, A. F., Desimone, R., & Ungerleider, L. G. (2001). Contextual modulation in primary visual cortex of macaques. *Journal of Neuroscience, 21,* 1698–1709.

Schira, M. M., Fahle, M., Donner, T. H., Kraft, A., & Brandt, S. A. (2004). Differential contribution of early visual areas to the perceptual process of contour processing. *Journal of Neurophysiology, 91,* 1716–1721.

Scholte, H. S., Witteveen, S. C., Spekreijse, H., & Lamme, V. A. F. (2006). The influence of inattention on the neural correlates of scene segmentation. *Brain Research, 1076,* 106–115.

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.

Sillito, A. M., Grieve, K. L., Jones, H. E., Cudeiro, J., & Davis, J. (1995). Visual cortical mechanisms detecting focal orientation discontinuities. *Nature, 378,* 492–496.

Super, H., Spekreijse, H., & Lamme, V. A. F. (2001). Two distinct modes of sensory processing observed in monkey primary visual cortex (V1). *Nature Neuroscience, 4,* 304–310.

Zipser, K., Lamme, V. A. F., & Schiller, P. H. (1996). Contextual modulation in primary visual cortex. *Journal of Neuroscience, 16,* 7376–7389.