Differential Contribution of Early Visual Areas to the Perceptual Process of Contour Processing

Mark M. Schira,1,2 Manfred Fahle,2 Tobias H. Donner,1 Antje Kraft,1 and Stephan A. Brandt1
1Department of Neurology, Charité Humboldt University of Berlin, 10117 Berlin; and 2Department of Human Neurobiology, University of Bremen, 28359 Bremen, Germany

Submitted 16 April 2003; accepted in final form 20 November 2003

Schira, Mark M., Manfred Fahle, Tobias H. Donner, Antje Kraft, and Stephan A. Brandt. Differential contribution of early visual areas to the perceptual process of contour processing. J Neurophysiol 91: 1716–1721, 2004. First published December 10, 2003; 10.1152/jn.00380.2003. We investigated contour processing and figure–ground detection within human retinotopic areas using event-related functional magnetic resonance imaging (fMRI) in 6 healthy and naive subjects. A figure (6° side length) was created by a 2nd-order texture contour. An independent and demanding foveal letter-discrimination task prevented subjects from noticing this more peripheral contour stimulus. The contour subdivided our stimulus into a figure and a ground. Using localizers and retinotopic mapping stimuli we were able to subdivide each early visual area into 3 eccentricity regions corresponding to 1) the central figure, 2) the area along the contour, and 3) the background. In these subregions we investigated the hemodynamic responses to our stimuli and compared responses with or without the contour defining the figure. No contour-related blood oxygenation level–dependent modulation in early visual areas V1, V3, VP, and MT+ was found. Significant signal modulation in the contour subregions of V2v, V2d, V3a, and LO occurred. This activation pattern was different from comparable studies, which might be attributable to the letter-discrimination task reducing confounding attentional modulation. In V3a, but not in any other retinotopic area, signal modulation corresponding to the central figure could be detected. Such contextual modulation will be discussed in light of the recurrent processing hypothesis and the role of visual awareness.

INTRODUCTION

Detection of contours is an essential early step in visual perception, and has been extensively studied with psychophysical, electrophysiological, and imaging techniques. Previous functional magnetic resonance imaging (fMRI) studies in humans have shown that perception of contours leads to modulation of the blood oxygenation level–dependent (BOLD) signal in early visual areas respecting their retinotopic organization (Mendola et al. 1999; Reppas et al. 1997; Skiera et al. 2000). These studies used passive viewing tasks, where subjects fixate a central cross, while figures containing real or illusory contours are presented. Although real contours led to significant modulations of the BOLD response in all retinotopic areas from V1 through V4, two types of illusory contours differed in the strength of activation of retinotopic areas including V3a, V4v, V7, and V8 bilaterally (Mendola et al. 1999). However, passive viewing tasks suffer from critical drawbacks. Huk et al. (2001) demonstrated that fMRI experiments using visual stimuli are very sensitive to attentional modulation. Differentiation of attentional modulation and modulation attributed to mostly stimulus-driven perceptual processes is critical for the interpretation of functional imaging data related to contour processing. Therefore attentional control is needed to clearly separate perceptual processes from neuronal activity reflecting attentional modulation.

Kastner et al. (2000) used attentional control and texture-defined contour stimuli; however, their study did not discriminate between signal modulations at different eccentricities within the relevant visual areas, which is necessary to discriminate contour processing from figure detection. Although those parts of the cortex retinotopically corresponding to the contour should process this contour, any activity located within the central part (i.e., the figure) signals the presence of this figure or object. The work of Lamme and colleagues (Lamme et al. 1995, 1998; Supér et al. 2001) clarifies this point. They found contextual responses in primary visual cortex (V1) neurons, that is, an increased neuronal spiking rate related to stimuli containing a closed figure. Not only did the neurons with receptive fields along the contour respond, but also neurons with receptive fields inside the figure did, thus, representing a contextual response. According to these studies, V1 neurons respond to both stimuli within the receptive field and to the context. Furthermore, Supér et al. (2001) showed that this contextual response disappears if the monkey was not aware of the figure. These findings suggested a clear extension of the receptive fields of V1 neurons beyond their function as simple local filters and led to the recurrent processing hypothesis (Lamme and Roelfsema 2000; Roelfsema et al. 2002). In short, they speculate that the local receptive field properties determine a fast and early feed-forward sweep of activity, whereas the contextual response extending the receptive field results from a more complex and slower recurrent feedback process.

The aim of the present study was to reexamine the modulation of retinotopic subparts corresponding to the figure, the contour, and the background, while controlling for the particular attentional effort. An event-related design was used and thus, our experimental design resembles those used by Lamme and colleagues. Our subjects performed a demanding foveal letter-discrimination task, similar to that used by Braun (1994). At the same time, we presented contour stimuli without relevance to the letter task. In fact, subjects remained unaware that contour stimuli had been presented at all. By avoiding recurrent feedback, this design aimed at maximizing its sensitivity to reveal early steps of visual processing.
METHODS

Stimuli

SUBJECTS. Four male and two female students (21 to 29 yr, mean = 25 yr, SD = 3.1 yr) from the Humboldt University of Berlin served as subjects in the study, which was conducted in conformity with the Declaration of Helsinki. All subjects had no history of neurological or psychiatric disorders and reported normal visual acuity. Subjects were paid for their participation and signed a consent form.

CONTOUR STIMULI. To measure activity related to contour processing we designed a stimulus [rotator-stimulus (R-Stim)] consisting of 13 × 13 small rotating elements (0.6° diameter each). Each rotator consisted of 3 small dots arranged in a line (Fig. 1A). The 2 outer dots rotated around the central dot, thus creating the percept of a rotating line, without overall changes in luminance or contrast. This stimulus avoids luminance artifacts on a digital screen and is therefore well suited for fMRI experiments using an LCD projector. In half of the trials the 5 × 5 central elements rotated with a phase lag of 180 ms. Because of the phase lag, the inner rotators had a tilt relative to the outer rotators, thus creating the percept of an isoeccentric (3°) square (Fig. 1B). Because all rotators were moving in the same way it was the relative tilt that defined the square’s contour, not motion cues. This was confirmed in psychophysical pilot experiments before the fMRI study. We could show that the stimulus provided a clear figure–ground contrast even in an isoluminant red–green condition, whereas the contour percept disappeared with increasing speed of the rotators. If the rotators had contained a significant motion-related contour, then the figure–ground contrast should have been enhanced.

In the fMRI study we tested two conditions: first, the control condition without the square’s contour where the 13 × 13 rotators created a homogeneous array of 8° side length. Second, the figure condition with the additional contour (3° eccentricity on the horizontal meridian) was present. This contour separated the array into a figure and a background (Fig. 1B). The difference between the 2 conditions is well defined and easily detectable for an attending observer.

The two conditions were presented in a randomized rapid event-related design, consisting of two scanning sessions of 695 volumes each corresponding to 25-min duration. Each stimulus was presented for 800 ms and the interstimulus interval varied from 4 to 14 s (7.9 s average), resulting in 2 × 180 trials.

LETTER-DISCRIMINATION TASK. The subjects’ attention was directed centrally by performing a foveal letter-discrimination task (Fig. 1A). Subjects had to decide in a 2 alternative forced-choice (2AFC) paradigm, whether these 5 elements were all identical (i.e., 5 L’s or 5 T’s), or whether one differed (e.g., 4 T’s and 1 L). We chose this task because it has been studied extensively in former studies (Braun et al. 1994, 1998; Li et al. 2002). Subjects were instructed to perform the letter task as accurately as possible without wasting time. This task prevented subjects from consciously perceiving the square’s contour.

After the experiment subjects were informed about the presence of a contour in some of the stimuli. Each subject was immediately able to detect the contour, but none had realized that the contour had also been present during the experiments. Before the fMRI sessions, subjects participated in a training session to become accustomed to the letter-discrimination task.

MAPPING STIMULI. Recall that the control condition created a homogeneous array. In the figure condition, however, this array was subdivided in 3 parts: the parafoveal figure, the contour of the square, and the background. To analyze the data of the R-Stim recordings, we accordingly divided each of the early visual areas (V1–V4) into three subregions [regions of interest (ROIs)]. The inner figure region is represented by the parafoveal space between 1 and 2.5° eccentricity, the neighboring contour area ranged from 2.5 to 3.5°, and the outer peripheral ROI ranged from 3.5 to 8°. To separate these three eccentricity ranges we designed a customized eccentricity-mapping experiment (Fig. 1, C and D), which provided localizers for the ROIs.

To define the borders separating early visual areas we combined classical checkerboard stimuli and sequences from popular television comic strips (Sereno et al. 2001). All mapping stimuli contained a central fixation dot and subjects were instructed to maintain fixation.

Equipment and fMRI protocols

The visual stimuli were displayed by ERTS (BeriSoft Cooperation, Frankfurt, Germany) and projected by an NEC LCD projector and a custom-made lens on a small back-projection screen (Daplex, 20 × 15 cm) mounted in front of the head coil. Subjects viewed the screen by way of a mirror at a distance of about 24 cm. MRI data were acquired using a 1.5-T Magnetom Vision (Siemens Medical Systems, Erlangen, Germany) equipped with an EPI booster and a standard head coil. Functional measurements were performed using single-shot EPI sequences (TE = 60 ms; FA = 90°). Structural 3-D data sets were acquired in the same session, using a T1-weighted sagittal MP-Rage sequence (TR/TE = 10/4 ms; FA = 12°; TI = 100 ms; voxel size = 1 mm³).
High-quality 3-D data sets for each subject were previously recorded using a T1-weighted sagittal FLASH sequence (TR/TE = 38/5 ms; FA = 30°; voxel size = 1 mm$^3$) with 2 acquisitions for excellent gray–white contrast for accurate segmentation and reconstruction of individual surface structures.

Data analysis

fMRI data were analyzed using the BrainVoyager 4.6 software package (BrainInnovation, Maastricht, The Netherlands). Functional data preprocessing contained motion correction, linear trend removal, and slight Gaussian 3D spatial smoothing [3 mm full width at half maximum (FWHM) for meridian mapping or 4 mm for eccentricity mapping and figure–ground measurements]. The block-design measurements (meridian and eccentricity mapping) were also temporally smoothed (4-s FWHM). The event-related figure–ground measurements were slice scan time corrected.

We segmented and reconstructed the surface of the white matter and produced flat maps from the high-resolution structural MRI images of each subject. Based on the results of the mapping experiments as described above ROIs were defined. Each of the early visual areas (V1, V2d, V2v, Vp, V3, V4v, V3a) contained 3 eccentricity-based ROIs: parafoveal, contour, and periphery (Fig. 2d). Further ROIs, which were identified in only 4 of the 6 subjects covered area MT+ and area LO. MT+ was localized with a low-contrast motion stimulus similar to the one described by Tootell et al. (1997). The ROI covering LO (Grill-Spector et al. 1999) was defined as the region located between the ventral border of MT+ and the foveal representation of V4v and V3. LO was thus defined by the location of the neighboring areas and not by specific functional characteristics. Thus definition of ROIs was not confounded with data from figure–ground experiments.

To test whether the responses of the 2 conditions of the R-Stim (figure vs. control) differ significantly, we performed paired $t$-testing on the hemodynamic response curves for all 23 ROIs. We defined the significance level at $P < 0.05$ after Bonferroni correction.

RESULTS

The mapping measurements allowed us to define 23 ROIs (Table 1). Each of the selected ROIs showed a clear hemodynamic response ranging from 0.3 to 0.7% signal change in both conditions (Fig. 2, Table 2). In all ROIs of V1, however, the response to the stimulus was equally strong, irrespective of the presence or absence of contours in the stimulus, implying no contour-related signal in this area (Fig. 2, a–c). In both ventral and dorsal V2, the response was stronger for the stimuli with contour than for the control (Fig. 2, e and f), whereas signals of the parafoveal and peripheral ROIs of V2 remained un-
TABLE 1. Size of regions of interest (ROIs)

<table>
<thead>
<tr>
<th>ROI</th>
<th>V1</th>
<th>V2v</th>
<th>V2d</th>
<th>V3</th>
<th>Vp</th>
<th>V3a</th>
<th>V4v</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parafoveal</td>
<td>445 ± 202</td>
<td>230 ± 63</td>
<td>291 ± 107</td>
<td>295 ± 86</td>
<td>337 ± 221</td>
<td>224 ± 77</td>
<td>236 ± 168</td>
</tr>
<tr>
<td>Contour</td>
<td>415 ± 239</td>
<td>290 ± 152</td>
<td>283 ± 171</td>
<td>326 ± 117</td>
<td>266 ± 174</td>
<td>380 ± 129</td>
<td>261 ± 188</td>
</tr>
<tr>
<td>Periphery</td>
<td>447 ± 190</td>
<td>295 ± 110</td>
<td>331 ± 106</td>
<td>388 ± 200</td>
<td>305 ± 121</td>
<td>337 ± 143</td>
<td>315 ± 193</td>
</tr>
</tbody>
</table>

Values are means ± SD, averaged over subjects and hemispheres. Because one functional voxel was approximately 36 mm³, the 445 mm³ value of the parafoveal ROI in V1 (and therefore the results in Table 2 and Fig. 2a) reflects about 12 functional voxels.

changed (Table 2). In higher visual areas, we found a strong contour-related signal modulation in LO, whereas in MT+ no modulation could be detected (Fig. 2, g and h).

Table 2 summarizes the responses of all the ROIs (except LO and MT+: see Fig. 2, g and h for their results), showing the peak of the hemodynamic response to the R-Stim. The response to stimuli containing the figure increases significantly (marked bold) in several ROIs. This always occurred (except in area V3a) in ROIs corresponding to the eccentricity at which the figure–ground boundary was located (= ROI: contour). V3 and VP showed no or only weak (not significant) contour-related modulation of the response. The areas V3a and V4v showed a significant signal increase in the contour ROIs. In V3a, this increase was present not only in the contour ROI but also in the parafoveal ROI, which suggests a response to the area inside the square (i.e., a figure response).

DISCUSSION

Applying a foveal attention-demanding letter-discrimination task prevented subjects from noticing the peripheral contours and avoided contour-related attentional modulations. Thus signal modulations should be related primarily to preattentive contour processing. We demonstrated that it is possible to measure a significant signal modulation related to a processing of locally invisible contours, even after exclusion of attentional modulation. We found no contour-related modulation in V1, V3, VP, and MT+, but significant modulation in V2, V3a, V4v, and LO.

Differences from previous studies

A prominent difference between our results and those of comparable fMRI studies (Kastner et al. 2000; Mendola et al. 1999; Reppas et al. 1997; Skiera et al. 2000) is the pronounced signal modulation in V2, whereas in V1 and V3/Vp modulation could not be detected. Of course, it cannot be concluded that there was a complete absence of contour-related modulations in V1. Nevertheless, we emphasize that the hemodynamic responses in V1 to both conditions of the R-Stim (reflecting the response related to the receptive field properties) were almost identical. We believe the different results in V1 and V2 reflect the fact that the processing of the square’s contour modulates V2, but not V1, whereas additional attentional modulation is prevented by the foveal letter-discrimination task.

Response in the figure part

Why was significant signal modulation in the figure part absent in V1 and V2? Such figure response was described by Lamme and colleagues (1995, 1998), who found modulation of firing rates of V1 neurons whose receptive fields covered exclusively the figure part of a figure–ground stimulus. A possible explanation for the lack of significant fMRI activity is supplied by the “recurrent processing” hypothesis (Lamme and Roelfsema 2000; Roelfsema et al. 2002; Supér et al. 2001). This hypothesis suggests that conscious figure perception is necessary to induce a figure response of this kind. By contrast, our stimulus was designed to avoid conscious figure perception and thus, did not induce a signal modulation within the figure part. In this context one might speculate that the absence of significant V1 modulation in our experiments reflects the fact that subjects were not aware of the presence of the figure (Supér et al. 2001). In summary V1 should be modulated by contour stimuli either if 1) neurons contribute to the processing of the global contour (bottom up), 2) if attention is attracted (top down), or 3) by conscious perception of the contour (recurrent processing). We hypothesize that those premises were not given in our experimental paradigm, although the last two points are difficult to prove.

TABLE 2. Peak responses of all ROIs placed in early retinotopic areas from V1 to V4v

<table>
<thead>
<tr>
<th>Visual Area</th>
<th>Parafoveal</th>
<th>Contour</th>
<th>Periphery</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>Figure</td>
<td>Control</td>
<td>Figure</td>
</tr>
<tr>
<td>V1</td>
<td>0.35 ± 0.035</td>
<td>0.36 ± 0.035</td>
<td>0.36 ± 0.035</td>
</tr>
<tr>
<td>V2v</td>
<td>0.26 ± 0.035</td>
<td>0.30 ± 0.034</td>
<td>0.27 ± 0.037</td>
</tr>
<tr>
<td>V2d</td>
<td>0.22 ± 0.033</td>
<td>0.25 ± 0.032</td>
<td>0.29 ± 0.038</td>
</tr>
<tr>
<td>V3</td>
<td>0.23 ± 0.031</td>
<td>0.24 ± 0.031</td>
<td>0.23 ± 0.031</td>
</tr>
<tr>
<td>Vp</td>
<td>0.20 ± 0.044</td>
<td>0.23 ± 0.043</td>
<td>0.19 ± 0.029</td>
</tr>
<tr>
<td>V3a</td>
<td>0.16 ± 0.03</td>
<td>0.23 ± 0.03</td>
<td>0.19 ± 0.037</td>
</tr>
<tr>
<td>V4v</td>
<td>0.16 ± 0.039</td>
<td>0.17 ± 0.038</td>
<td>0.16 ± 0.048</td>
</tr>
</tbody>
</table>

Values are percentage signal change ± SE. ROIs in which the responses of 2 conditions differed significantly (P < 0.05) are in bold.
Response differences of V1 and V2

Why does V2 but not V1 respond to the contour stimulus? First, the receptive fields of neurons in V2 are significantly larger than those in V1 (Smith et al. 2001). A local receptive field–based receptor must cover at least 2 rotators to detect the contour. Exact receptive field sizes are unknown for humans, but in the macaque the receptive field sizes at an eccentricity of 3° (position of the square’s contour) range in V1 ≤ 0.8° and in V2 from 1.2 to 2.0° (Lee et al. 2001). Thus, the receptive fields of 2° would be able to span two rotators, whereas 0.8° should be too small because two rotators span 1.4°.

Second, Hegde and Van Essen (2001) found V2 neurons with selectivity to rather complex texture elements like arcs or circles. These neurons seem to prefer complex texture properties, such as those responsible for contour segregation of our stimulus. So far, cells with comparable receptive field properties could not be identified in V1. These cells seem to be well suited to detect the relative tilt of our rotators. Thus the receptive field size alone probably does not explain the absence of significant V1 modulation. In fact, von der Heydt and Peterhans (1989) compared the response of monkey V1 and V2 neurons to abutting line gratings and other stimuli creating an illusory or anomalous contour percept. They found a significant response to their stimuli in V2, but not in V1, neurons. This finding was recently supported and extended by Ramsden et al. (2001) using optical imaging.

V3 and V3a

The results for areas V3 and V3a appear to be in line with the physiology of these areas in the macaque monkey. Unfortunately, the homology between macaques and humans is problematic especially in V3a (Vanduffel et al. 2001). It is important to note that, even though our stimulus contains motion, this information did not produce the square’s contour. V3 showed a BOLD response for the homogeneous R-Stim but this response was not further enhanced by the square’s contour. In contrast, V3a neurons showed clear contour-related modulation. The V3a response is particularly strong and also remarkable because V3a is the only area with significant modulation in the figure part. V3a neurons have relatively large receptive fields and they contain a complete and contiguous representation of the visual field. These properties are good requisites for figure integration, but they do not satisfactorily explain why exclusively V3a contained contextual responses.

The role of V4v

V4v is known to play a central role in attentional processes relevant for some types of visual search. It is supposed to play a central role in performing attentional shifts, for example, during visual conjunction search. The letter-discrimination task was designed by Braun (1994) to prevent these attentional shifts. Hanazawa and Komatsu (2001) demonstrated selectivity of V4 neurons in macaques for complex texture properties. Because we controlled attentional modulation, we argue that the V4v modulation expresses contour processing alone.

LO and MT+

Although the response of MT+ remains unaffected by the square’s contour, LO shows a clear modulation indicating the processing of the contour. This activation is possibly attributable either to a feed-forward flow of contour information derived from V2 or to a feature-contour system detecting salient regions as postulated by Grossberg et al. (1994). Recent fMRI results (Stanley and Rubin 2003) propose that LO plays a critical role in a feature-contour system.

In conclusion, contour processing was investigated with a 2nd-order contour stimulus using an event-related fMRI design at a high spatial resolution. Control of attentional modulation during contour perception prevented unspecific BOLD responses. We suggest that the contour stimulus was primarily processed in V2 and then feed-forward propagated to V3a, V4v, and LO. Interestingly, neither V1 nor V2 showed contextual modulation, but V3a did. The combination of eccentricity-matched fMRI responses with an attentional control paradigm leads to activation patterns different from those described in similar studies. We speculate that we were imaging primarily feed-forward processing of contours within one processing stream.

References


Lamme VA, Zipser K, and Spekreijse H. Figure-ground activity in primary visual cortex is suppressed by anesthesia. Proc Natl Acad Sci USA 95: 3263–3268, 1998.


J Neurophysiol • VOL 91 • APRIL 2004 • WWW.JN.ORG
Roelfsema PR, Lamme VA, Spekreijse H, and Bosch H. Figure-ground segregation in a recurrent network architecture. J Cogn Neurosci 14: 525–537, 2000.


