Perceptual segmentation and sustained activity in primary visual cortex

<table>
<thead>
<tr>
<th><strong>Journal:</strong></th>
<th><em>Journal of Vision</em></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Manuscript ID:</strong></td>
<td>JOV-00760-2007</td>
</tr>
<tr>
<td><strong>Manuscript Type:</strong></td>
<td>Special Issue</td>
</tr>
<tr>
<td><strong>Date Submitted by the Author:</strong></td>
<td>01-Oct-2007</td>
</tr>
<tr>
<td><strong>Complete List of Authors:</strong></td>
<td>Teichert, Tobias; University Marburg, Physics Eckhorn, Reinhard; University Marburg, Physics Wachtler, Thomas; University Marburg, Physics</td>
</tr>
<tr>
<td><strong>Key Words/Topics:</strong></td>
<td>perceptual organization, 3D surface and shape perception, binocular vision</td>
</tr>
<tr>
<td><strong>Special Issue:</strong></td>
<td>Perceptual Organization and Neural Computation</td>
</tr>
</tbody>
</table>
Perceptual segmentation and sustained activity in primary visual cortex

Tobias Teichert*, Reinhard Eckhorn, Thomas Wachtler

Philipps University, Department of Physics, NeuroPhysics Group, D-35037 Marburg, Germany

Email: Tobias Teichert* - tobias.teichert@physik.uni-marburg.de; Reinhard Eckhorn - reinhard.eckhorn@physik.uni-marburg.de; Thomas Wachtler - thomas.wachtler@physik.uni-marburg.de;

*Corresponding author

Abstract

In several studies, enhancement of firing rates in macaque primary visual cortex (V1) has been found to covary with the perceptual segmentation of a scene into figure and background (figure-ground enhancement). We tested whether differences in the perceptual segmentation of a scene may cause firing rate enhancement in V1 when no local cues indicate the correct segmentation. To this aim we manipulated the perceptual segmentation of a scene consisting of two overlapping rectangles by means of horizontal disparity. In the presence of local disparity cues on the surfaces of the rectangles, neurons enhanced their activity when coding part of the foreground compared to the background object starting 100 ms after stimulus onset. This finding demonstrates that figure-ground enhancement occurs with complex scenes involving overlapping objects. However, figure-ground enhancement was not observed in the absence of local cues, despite the fact that the perceptual segmentation of the scene, as indicated by the animal's behavioral responses, was the same as with local cues. This suggests that under these conditions, the perceptual segmentation of a scene by itself is not sufficient to modulate firing rates in primary visual cortex.

Keywords: perceptual organization, 3D surface and shape perception, binocular vision
Introduction

The perceptual segmentation of a visual scene into distinct objects and background is one of the most important tasks of the visual system. In several studies, firing rates of neurons in primary visual cortex (V1) were found to correlate with the perceptual segmentation of a scene. The sustained activity of neurons coding a certain part of a scene which was set apart from the background by differences in a visual feature like, e.g., orientation, disparity, spatial phase or color, was enhanced compared to conditions in which the neurons coded an identical part of the background (figure-ground enhancement) [1-4]. Despite extensive research over the past years, it is not yet clear whether the effect is caused by the perceptual interpretation of the scene into object and background, or whether it is caused by the manipulations of the physical stimulus which were necessary to change the perceptual segmentation of the scene.

Some recent results support the notion that figure-ground enhancement is not caused by the perceptual segmentation of the scene. Zipser et al. [2] reported that the strength of the figure-ground enhancement depends on the size of the figural region. Their results suggest that for disks of a diameter of more than 8 degrees visual angle, no figure-ground enhancement can be observed. This range is comparable to other reports of spatial interactions in V1 at comparable eccentricities [5-8]. Also, results from Marcus et al. [3] show that the effect of figure-ground enhancement is stronger for orientation than for spatial phase. They argue that both findings can be taken as evidence against the assumption that figure-ground enhancement is caused by the perceptual interpretation of the scene alone, as a region may be segregated from the background independent of its size or whether orientation or spatial phase is the defining feature. Further, Rossi et al. [9] interpret their findings that figure-ground enhancement can also be observed for regions which are not completely segregated from the background as evidence against the assumption that response strength in V1 reflects the perceptual organization of a scene. The modulations of sustained responses in V1 might be caused by lateral interactions within V1 as suggested, for example, in the models of Li [10,11]. However, other findings can be interpreted as evidence in favor of the hypothesis that figure-ground enhancement is being caused by the perceptual segmentation of a scene. Zipser et al. [2] surrounded a rectangular region of a binocularly presented random dot background with a frame of either crossed or uncrossed disparity. This gave rise to the impression of either a frame floating in front of a background, or a trench which segregated the enclosed rectangular region from the background. Figure-ground enhancement was observed only in the latter condition. This finding suggests that figure-ground enhancement is not merely a result of feature contrast, and thus needs to be distinguished from the effects of surround suppression [5,8,12-18]. The authors even speculated that figure-ground enhancement might
Figure 1: **Disparity capture** describes the fact that parts of a scene lacking unique disparity cues, in this case the surfaces of the rectangles, are nevertheless perceived in the depths indicated by regions providing unique cues, in this case the vertical borders of the rectangles. (A) Monocular versions of the stimuli whose depth percept is to be manipulated. (B) Regions in which manipulations of disparity lead to differences between the left- and right-eye version of the picture are coded in the corresponding color. (C) Surfaces are colored corresponding to perceived depth. Note that small manipulations of disparity (B) lead to a pronounced difference in the perceptual segmentation of the scene (C).

The present study was aimed at examining whether figure-ground modulations of firing rate in V1 are caused by the perceptual segmentation of a visual scene. Figure-ground enhancement has so far mainly been investigated using stimuli which make it difficult to distinguish whether it is caused by the perceptual segmentation of the scene or by the differences in the physical stimulus used to manipulate the
segmentation of the scene. In order to make this distinction, we constructed stimuli for which very small manipulations in the periphery of the retinal image reliably altered the perceptual segmentation of the entire scene. This was achieved by manipulating relative disparity and exploiting an effect known as disparity capture. Disparity capture describes the effect that regions which lack unique disparity cues will be perceived in the same depth as the disparity of non-horizontal borders of these regions indicate (see Fig. 1 and Methods). Thus, the disparity cues used to alter the segmentation of the scene had to captured from regions outside the compound receptive fields of the neurons under investigation (see Fig. 2b and c).

In addition we used stimuli for which the disparity cues used to alter the segmentation of the scene were inside the compound receptive fields (see Fig. 2a). These conditions with or without unique local disparity information will be referred to as locally unambiguous and locally ambiguous, respectively.

Human observers with normal depth perception confirmed the effects of disparity on the perceptual segmentation of the scene in both conditions. A task associated with the correct segmentation of the scene was performed equally well by humans and the monkey, indicating that the monkeys perceptual segmentation of the scene was manipulated in an identical fashion (see also [23]).
In the present set of experiments we examined the occurrence of figure-ground enhancement in visual scenes comprising two overlapping objects. The locally unambiguous condition can be considered a test of the external validity of figure-ground enhancement under these conditions. The locally ambiguous condition tests whether figure-ground enhancement may be caused by perceptual segmentation of the scene even in the absence of local cues.

**Methods**

*Experimental setup*

Stimuli were presented binocularly on a Wheatstone stereoscope. A mirror system was placed in front of the monkey’s head such that left and right eye could independently be stimulated by two 15° monitors located 100 cm left and right of the monkey’s head. In this setup, one pixel corresponded to approximately 1.2 minutes of arc (for more details on the setup see [24]).

*Stimuli*

Stimuli consisted of two overlapping rectangles (6.5°×2° and 2°×6.5°, respectively) monocularly forming a cross (see Fig. 2). The relative disparity of the two rectangles was manipulated to mimic stimulus configurations with either the horizontal or the vertical rectangle in front (see Fig. 1). The lateral position of the horizontal rectangle varied in the same range as the relative disparity offset. Thus, no monocular cues regarding the segmentation of the scene were available in any case.

Surface structure of the rectangles was manipulated to yield three different conditions. For the *random-dot stimuli* each pixel on the surface was randomly chosen to be either black or white. Consequently, valid disparity cues were available on the entire surface. These stimuli will be referred to as *locally unambiguous*. Surfaces of the *uniform stimuli* were of a uniform white. For these stimuli, valid disparity cues were only present on the vertical borders of the rectangles. These stimuli will be referred to as *locally ambiguous*.

The surfaces of the *cyclic-dot stimuli* were covered with a cyclic pattern (3×4 black pixels centered on 6×6 white pixels, see Figure 2). The spatial period of the pattern (6 pixels) was chosen identical to the disparity offset of between the two rectangles (Wallpaper illusion). Like the uniform surfaces, these stimuli have no unique disparity cues on their surfaces. These stimuli will be referred to as *locally ambiguous* as well.
Experiment I

In Experiment I the uniform and the random-dot stimuli were presented. During some recording sessions both stimulus types were presented randomly interleaved, while during others only one stimulus type was presented. The relative disparity between the two rectangles was manipulated exclusively by changing the absolute horizontal disparity of the horizontal rectangle. The vertical rectangle remained in the plane of fixation at all times. Further, the vertical rectangle extended either to the right or the left of the horizontal midline of the horizontal rectangle (Fig. 2).

Experiment II

In Experiment II the cyclic-dot and the random-dot stimuli were presented randomly interleaved. Absolute disparity of either the horizontal or the vertical rectangle was manipulated while the other rectangle remained in the plane of fixation. Manipulating absolute disparity of both rectangles lead to a doubling of the number of stimulus conditions. To keep the number of repetitions per condition constant, the number of conditions was decreased accordingly by fixing the position of the vertical rectangle relative to the midline.

Task and procedure

Task of the animal was to report which of the two rectangles appeared in front. The monkey initiated each trial by touching a lever. A fixation spot appeared and after a fixation period of variable length (600 to 1000 ms) a stimulus was presented for 1100 ms. After another variable fixation period (300 to 800 ms) the fixation spot dimmed. This cued the monkey to report whether he had perceived the vertical or the horizontal rectangle in front by moving a lever up or down, respectively. Responses were considered valid if the monkey held on to the lever until the fixation spot dimmed and responded within the next 1000 ms. During the whole trial, eye position was restricted to a fixation window of .5°.

Preparation and Recordings

Experiments were performed with a male macaque monkey. Preparation and recording were in accordance with German laws of animal maintenance and experimentation and the guidelines published in the NIH Guide for the Care and Use of Laboratory Animals. Extracellular electrical activity was recorded from the upper layers of primary visual cortex, using a 2×8 array of singly moveable quartz-isolated platinum-tungsten fiber-microelectrodes [25]. Raw signals (1 Hz - 10 kHz) from each electrode were filtered online to obtain multi-unit activity (MUA: 1-10 kHz band-passed, full-wave rectified, 140 Hz low-passed,
sampled at 500 Hz) and local field potentials (LFP: 250 Hz low-passed, sampled at 500 Hz). During 20 recording sessions a total of 320 microelectrode-penetrations were made. Movements of the left eye were monitored with an infrared camera system (Thomas Recording, Giessen) and sampled at a rate of 500 Hertz.

**Classification of recording sites**

Population RFs were mapped using small bright spots with a Gaussian luminance distribution ($\sigma = 4.9$ minutes of arc) flashed for 50 ms at randomly chosen positions on a regular 16×16 grid [26]. MUA amplitude as a function of spot position was interpolated and contour lines at a level of 70% of the maximum response strength were calculated. The corresponding area in visual space was defined as the receptive field (RF) of the multiple unit recorded at this site. Typical receptive field sizes were in the range of 0.25 to 0.5 degrees.

Recordings from the 2×8 microelectrodes gave rise to a 2×8 array of RFs. The microelectrode-array was positioned on the cortex in a way that the long axis of the RF-array was approximately horizontal in visual space. Stimuli were located relative to the fixation point such that half of the receptive fields were right and left of the midline of the horizontal rectangle, respectively. All receptive fields were at least 2 degrees from the vertical borders of the horizontal rectangle. Recording sites were classified as either left (L), right (R) or on the border (B) according to the position of their RFs relative to the midline.

**Data processing and statistical analysis**

Stimulus-evoked multi-unit activity (MUA) and local field potentials (LFP) were obtained by averaging all responses to identical stimulus configurations aligned to stimulus onset. Evoked potentials and corresponding standard errors were low-pass filtered with a cutoff frequency of 30 Hz for illustration purposes. Response onset latencies (OL) were defined as the first time when the response strength exceeded a threshold, which was defined as mean response strength on the baseline plus 4 times the mean standard error for 3 consecutive time steps. If no onset was detected within 80 ms after stimulus onset, no onset latency was defined.

Welch-tests ($\alpha=0.05$, two-sided) for differences in the condition with the vertical compared to the horizontal rectangle in front were performed for mean response strength in sliding rectangular time windows with 4, 8, 16, and 32 ms length. Tests were conducted for 200 consecutive time periods from 30 ms to 428 ms after stimulus onset. A time point was considered to show significant differences when (1) tests of all 4 window
Figure 3: **Evoked activity of a typical recording site.** LFP (A) and MUA (B) evoked by (a) random-dot (b) and uniform stimuli. Blue and green lines correspond to conditions with the horizontal and vertical rectangle in front, respectively. Time periods with significant differences between the conditions are marked with yellow (see methods for details). Evoked activity was weakest when the receptive field was on the horizontal rectangle (I) and the horizontal was behind the vertical rectangle (green line).

lengths were significant and (2) tests with 4 ms window length of the two following periods were significant. This method yielded conservative estimates of significant differences between the two conditions.

Apart from the sliding Welch tests, all statistical analyses were performed with the statistics package R version 2.4.0 [27]. Non-parametric tests (two-sample Wilcoxon test and the Spearman rank-correlation test) were used throughout.

**Results**

During training sessions, psychometric functions for the detection of relative disparity differences were measured. During recording sessions, disparity values well above the monkey’s threshold were used.

**Experiment I**

Averaged over all 10 recording sessions, performance was at 97% ± 2% (only trials with correct fixation and valid response). Performance for random-dot rectangles was slightly but significantly better than performance for rectangles with uniform surfaces (99% vs. 96%, binomial test over all trials). Further, responses were more accurate when the vertical rectangle was to the right of the midline of the horizontal
Figure 4: **Figure-ground enhancement.** Percentage of recording sites with significantly stronger LFP (A) and MUA (B) in the condition with the horizontal (blue) and the vertical (green) rectangle in front. Results are presented separately for conditions with the RFs on the overlapping part of the rectangles (I) or on the branch of the horizontal rectangles (II). Top row: For the random-dot stimuli, responses of a large proportion of recording sites depended on the depth order of the rectangles. For RFs on the region of overlap of the rectangles, responses were stronger when the vertical rectangle was in front (green line). For RFs on the non-overlapping part of the horizontal rectangle, responses were stronger when the horizontal rectangle was in front (blue line). (b) No such modulations were observed for the uniform rectangles.

Neuronal responses in V1

Figure 3 shows the evoked activity of a typical recording site. Both MUA and LFP indicate stronger activity for the random-dot than for the uniform surfaces. This difference was especially pronounced for the multi-unit signal (see Fig. 3B). For the random-dot stimuli, in conditions when the RF was located on the horizontal rectangle outside the region of overlap (see Fig. 3II, upper row), responses were stronger when the horizontal rectangle was in front (green line). This pattern can be described as stronger responses for neurons coding the unoccluded rectangle. Prevalence and timing of differences in response strength between the conditions with the horizontal compared to the vertical rectangle in front was further quantified at the population level. To this aim we calculated the percentage of recording sites which exhibited significant differences in response strength as a function of time from stimulus onset (see Fig. 4). Green and blue lines indicate the proportion of units with stronger activity in conditions with the vertical rectangle compared to when it was to the left (98% vs. 96%, p =0.0029, binomial test).
Figure 5: Mean evoked MUA to random-dot (a) and cyclic-dot stimuli (b) of all recording sites recorded in experiment 2. The green and blue lines correspond to conditions with the vertical and horizontal rectangle in front, respectively. In (A) segmentation of the scene was manipulated via disparity of the vertical rectangle, in (B) via disparity of the horizontal rectangle. Response strength for the unambiguous random-dot stimuli (top row) and for the ambiguous cyclic-dot stimuli (bottom row) were similar. However, response modulations depending on the depth order of the rectangles occur with the locally unambiguous but not with the locally ambiguous stimuli.

and horizontal rectangle in front, respectively.
For the uniform surfaces only few recording sites showed significant modulation of activity. For the random-dot rectangles, significant modulations of response strength were common, as illustrated in Figure 3. Recording sites with RFs on the non-overlapping part of the horizontal rectangle typically responded stronger when the horizontal rectangle was in front. This increase in response strength starts about 100 ms after stimulus onset. Significant modulations of response strength were less frequent for recording sites with their RFs on the overlapping part of the rectangles (LFP: 50% and MUA: 20%, see Fig. 4I). More recording sites responded stronger when the vertical than when the horizontal rectangle was in front.

Eye movements
Changes in mean eye position were detectable around 200 ms after stimulus onset (data not shown). For all stimuli we observed a systematic deviation of eye position to the upper right. The deviation was in the order of 0.3 degree visual angle. The upward component was not critical as it did not change the relative
Figure 6: **Figure-ground enhancement.** Percentage of recording sites with significantly stronger LFP (A) and MUA (B) in conditions with the horizontal (blue) and vertical rectangle (green) in front. Layout and symbols as in Fig. 4. Results for the conditions with the horizontal or the vertical rectangle in the plane of fixation are pooled. Results are comparable to experiment I. If the receptive fields were located on the branch of the horizontal rectangle, evoked activity was stronger when the horizontal rectangle was in front.

position of the RFs with respect to the midline. The rightward deviation, however, changed the relative position and consequently lead to the erroneous classification of some RFs (see Methods). Such misclassification might have lead to an attenuation of differences between conditions with RFs on the overlapping part of the rectangles and the branch of the horizontal rectangle. However, the clearcut distinction of response properties between these conditions (see Fig. 4), can be taken as evidence that these misclassifications did not affect our results to a large degree.

Microsaccades were observed during fixation. For the random-dot surfaces these fixational eye movements lead to a change of the stimulus in the receptive fields and consequently evoked synchronous bursts of activity from all recording sites. Microsaccades were rare in the first ~350 ms after stimulus onset. To avoid contamination of the data with effects of eye-movements, all trials with microsaccades between 0 to 350 ms relative to stimulus onset were excluded from analysis. In the remaining trials no difference in fixational eye-movements could be detected.
Experiment II

The responses evoked in the locally ambiguous condition, i.e., by the uniform stimuli, were clearly weaker than the ones evoked in the locally unambiguous condition. To rule out that the lack of a modulation of response strength in the ambiguous condition was due to overall lower levels of activation, the cyclic-dot stimuli replaced the uniform stimuli in the ambiguous condition of experiment II. Due to the local luminance contrast, these stimuli should evoke stronger responses from neurons in V1 than the uniform surfaces.

The cyclic-dot stimuli evoked significant responses from all recording sites. Figure 5 shows the mean evoked activity. Response strength as well as response latency were comparable for random-dot and cyclic-dot stimuli. However, while for the random-dot stimuli response strength depended on whether the horizontal or vertical rectangle was in front, no such effect could be observed for the cyclic-dot stimuli (see Figure 5, upper versus lower row).

As in experiment I, prevalence and timing of differences in response strength between the conditions with the horizontal compared to the vertical rectangle in front were further quantified on the population level. The results corroborated those of experiment I. While for the random-dot stimuli response strength of most recording sites depended on which rectangle was in front, this was rarely the case for the cyclic-dot stimuli (see Fig. 6, upper versus lower row). For the random-dot stimuli, a large proportion of recording sites with RFs on the non-overlapping part of the horizontal rectangle enhanced their activity from 100 ms onward when the horizontal rectangle was in front (see Fig. 6IIA and 6IIB, upper rows).

In Experiment I, it was not possible to isolate effects of figure-enhancement from effects of selectivity to absolute disparity. In order to distinguish the influence of these two factors, the segmentation of the scene was manipulated by varying the disparity of either the vertical or the horizontal rectangle in experiment II. When the segmentation of the scene was manipulated by the disparity of the vertical rectangle, recording sites with RFs on the horizontal rectangle outside the region of overlap responded stronger when the horizontal rectangle was in front (see Fig. 6AII upper row), despite the fact that the retinal disparity of the random-dot pattern was identical in both conditions. In contrast, units with RFs on the overlapping part of the rectangles (see Fig. 6AI, upper row) tended to respond stronger when the vertical rectangle was in front, despite the fact that in both cases the RFs were on the foreground object. However, in these cases the absolute horizontal disparity of the random-dot pattern on which the RFs were located, differed between the two conditions (see Methods).
Discussion

Our study yielded two main findings. First, in the unambiguous case we observed enhancements of sustained responses of neurons coding the unoccluded rectangle. Independence of the effect from absolute disparity and the late onset around 100 ms after stimulus onset indicate that it can be considered an extension of figure-ground enhancement [1, 2] to the case of overlapping stimuli.

Second, while figure-ground enhancement was found in the presence of local cues (random-dot stimuli), it was not observed in conditions lacking local cues (uniform and cyclic-dot stimuli). This is especially noteworthy as the perceptual segmentation of the scene, as indicated by the responses of the monkey, was comparable in both conditions. Thus, in our setting, the perceptual segmentation of the scene does not seem to be sufficient to cause the figure-ground enhancement of neurons in V1. In the following we will discuss the two main findings in greater detail.

Neuronal responses in the unambiguous condition

Neuronal responses for the random-dot rectangles depended on relative disparity of the rectangles, i.e., the perceptual organization of the scene, and the position of the RF on the the branch or the overlapping part of the horizontal rectangle (see Fig. 4 and 6). Differences in response strength might be due to figure-ground enhancement or selectivity of neurons to absolute disparity which had to be varied in order to manipulate relative disparity between the rectangles. In experiment I the two possible sources were confounded, thus it was not possible to determine their relative contribution to the observed effects. Experiment II, however, showed that figure-ground enhancement as well as selectivity to absolute disparity contributed to the observed effects.

Figure-ground enhancement

Lamme et al. [1] found that neurons responded stronger when their RFs were located on a part of the scene that was segregated from the surrounding background by one of several possible cues like orientation, movement or disparity. As in our setting receptive fields were never located on the background but instead on one of two objects, the classical effect of figure-ground enhancement does not apply to the current situation. However, it can be extended in a straightforward manner to the situation of two partially overlapping objects. The extended version of figure-ground enhancement predicts stronger activity for neurons coding the non-occluded foreground object compared to the occluded background object.

Neurons with RFs on the overlapping part of the rectangles always coded the non-occluded object,
regardless of whether the vertical or horizontal rectangle was in front. In contrast, the receptive fields
located on the branch of the horizontal rectangle coded the non-occluded object only when the horizontal
rectangle was in front, and coded the occluded object otherwise. Thus, the observed pattern of response
strength can be explained by stronger responses to the non-occluded object. The late onset of this pattern
in the sustained phase, i.e., ~100 ms after stimulus onset is comparable to the late onset of figure-ground
enhancement reported in several other experiments [1, 2, 4].

While in experiment I it was not possible to separate the effects of figure-ground enhancement and
selectivity to absolute disparity, it was possible to do so in experiment II. In the conditions when the
segmentation of the scene was manipulated by changing absolute disparity of the vertical rectangle
(Fig. 6AII), absolute disparity of the horizontal rectangle was constant. Thus, the increase in the firing
rate of neurons located on the branch of the horizontal rectangle in conditions when they coded the
non-occluded object can not be explained by selectivity to absolute horizontal disparity. Instead, the effect
is compatible with the extended version of figure-ground enhancement.

**Disparity tuning**

Neurons in V1 are known to respond selectively to absolute horizontal disparity [29–35]. We varied the
absolute disparity of the rectangles to manipulate the segmentation of the scene. Thus, the differences in
response strength might to some extent reflect the selectivity of neurons to absolute disparity. To isolate
the effects of disparity from the effects of figure-ground enhancement, we analyzed response strength of
recording sites located on the overlapping part of the rectangles, which should be subject to figure-ground
enhancement in all conditions. In the second experiment we observed stronger responses for recording sites
when the disparity of the vertical rectangle was varied to manipulate the segmentation of the scene and the
vertical rectangle was in front (Fig. 6AIa). This observation is consistent with stronger responses for
crossed compared to zero disparity. However, the fact that no such effect was observed when the disparity
of the horizontal rectangle was manipulated (see Fig. 6BI, upper row) seems to contradict this notion. In
the following we propose a tentative explanation for this. When the horizontal rectangle was presented
with crossed disparity the monkey might have changed the vergence angle of his eyes to the depth-plane
defined by the disparity of the horizontal rectangle which was very close to the fixation point. This would
have resulted in identical absolute disparities for neurons coding the overlapping part of the rectangle
regardless of whether the horizontal rectangle was presented with crossed or uncrossed disparity. As we
recorded only the position of one eye it was not possible to verify this assumption. However, this would not
affect the interpretation of the results when the segmentation of the scene was manipulated by the absolute disparity of the vertical rectangle. As Bakin et al. [38] showed, vergence angle of a macaque monkey performing an identical fixation task as in our experiment is not affected by changes in disparity of a para-foveally presented stimulus.

**Neuronal responses in the ambiguous condition**

The ambiguous stimuli were used as a strict test of whether figure-ground enhancement might be caused by the perceptual segmentation of a scene. Neither modulations of firing rate associated with figure-enhancement nor modulations indicating selectivity to absolute horizontal disparity were found for the ambiguous stimuli. In the following we will discuss the lack of these effects.

**Figure-ground enhancement**

The data from the unambiguous stimuli suggest that the extension of figure-ground enhancement to the present setting which involves two overlapping objects is justified. If figure-ground enhancement were to be caused by the perceptual interpretation of a part of the scene as foreground, we should expect the same effects to be present for the ambiguous stimuli, as the perceptual segmentation of the scene, measured by performance in the relative disparity task, was comparable in both conditions. However, for the locally ambiguous stimuli no such effect was observed. This suggests that the mechanisms responsible for figure-ground enhancement in some way rely on local cues, possibly in the form of discontinuities in some feature map or a feature contrast between the surface of the figure and the background. EEG data from human subjects suggest that both factors might play a role [36,37].

**Disparity tuning**

None of the modulations in firing rate attributable to absolute disparity were observed for the ambiguous stimuli. This was not unexpected, because for the ambiguous stimuli no unique disparity cues were present in the RFs of the neurons. Any effects of absolute disparity would have to be mediated by integration of disparity information from outside the RFs. However, Bakin et al. [38] showed that while neurons in V2 do capture disparity from outside their RFs, neurons in V1 fail to do so. Our results are consistent with these findings.
Conclusion

Our results indicate that figure-ground enhancement, which has so far only been examined in conditions with a single object in front of a uniform background, can be extended to more complex viewing situations with two overlapping objects. These findings support the notion that enhancement of sustained activity in V1 might support perceptual grouping even in conditions which involving a challenging binding problem, as suggested previously [39].

The results from the locally ambiguous stimuli indicate that in the present setting, modulations of firing rate are not caused by the perceptual segmentation of a scene alone. Together with the results from Supér et al. [4] this suggests that the perceptual segmentation of the scene into figure and background might be a necessary but not a sufficient condition for figure-ground enhancement. Further experiments are needed to determine what kind of local cues are needed for figure-ground enhancement and how the two different sources of information are integrated. We conclude that to date it is not known whether figure-ground enhancement occurs during natural viewing, i.e., in conditions where the segregation of a figural region from the background might depend on the interaction of a variety of different visual features.

References


