

Coding of Border Ownership in Monkey Visual Cortex

Hong Zhou,¹ Howard S. Friedman,^{1,3} and Rüdiger von der Heydt^{1,2}

¹Krieger Mind/Brain Institute, ²Department of Neuroscience, and ³Department of Biomedical Engineering, Johns Hopkins University, Baltimore, Maryland 21218

Areas V1 and V2 of the visual cortex have traditionally been conceived as stages of local feature representations. We investigated whether neural responses carry information about how local features belong to objects. Single-cell activity was recorded in areas V1, V2, and V4 of awake behaving monkeys. Displays were used in which the same local feature (contrast edge or line) could be presented as part of different figures. For example, the same light–dark edge could be the left side of a dark square or the right side of a light square. Each display was also presented with reversed contrast.

We found significant modulation of responses as a function of the side of the figure in >50% of neurons of V2 and V4 and in 18% of neurons of the top layers of V1. Thus, besides the local contrast border information, neurons were found to encode the side to which the border belongs (“border ownership coding”). A majority of these neurons coded border ownership and the local polarity of luminance–chromaticity contrast. The others were insensitive to contrast polarity. Another 20% of the neurons of V2 and V4, and 48% of top layer V1, coded local contrast polarity,

but not border ownership. The border ownership-related response differences emerged soon (<25 msec) after the response onset. In V2 and V4, the differences were found to be nearly independent of figure size up to the limit set by the size of our display (21°). Displays that differed only far outside the conventional receptive field could produce markedly different responses. When tested with more complex displays in which figure–ground cues were varied, some neurons produced invariant border ownership signals, others failed to signal border ownership for some of the displays, but neurons that reversed signals were rare.

The influence of visual stimulation far from the receptive field center indicates mechanisms of global context integration. The short latencies and incomplete cue invariance suggest that the border-ownership effect is generated within the visual cortex rather than projected down from higher levels.

Key words: primate visual cortex; visual perception; perceptual organization; figure-ground segregation; awake macaque monkey; single-unit activity; nonclassical receptive fields; area V1; area V2; area V4

When neural function in the monkey visual cortex was first analyzed, it was concluded that the initial stages represent visual information in terms of local features, each neuron analyzing the small area of the retinal image covered by its receptive field, which occupies only a tiny fraction of the whole visual field (Hubel and Wiesel 1968, 1977). This notion has been modified by studies showing that responses evoked by a local stimulus can also be modulated by stimulation of a larger surround of that small area (which was then termed the “classical receptive field”; Nelson and Frost, 1978; Allman et al., 1985; Gilbert and Wiesel, 1990; Knierim and Van Essen, 1992; Pettet and Gilbert, 1992; Sillito and Jones, 1996). These findings have generally been interpreted as evidence for receptive field surrounds that either inhibit or facilitate the excitation generated by stimulation of the receptive field center. The surround influence might serve to enhance the sensitivity of the system for feature contrast, which could play a role in feature discrimination and visual search (Allman et al., 1985; Knierim and Van Essen, 1992) or it might serve to fill in visual scotomata (Pettet and Gilbert, 1992). Displays that produce the perception of illusory contours can also evoke responses when the actual stimulation is confined to areas outside the classical receptive field (von der Heydt et al., 1984; Peterhans and von der Heydt, 1989). These responses might be related to figure-ground mechanisms (von der Heydt et al., 1993; Baumann et al., 1997; Heitger et al., 1998).

Lamme et al. (Lamme, 1995; Zipser et al., 1996; Lee et al., 1998)

have recently discovered that responses of cells of V1 to textured stimuli are enhanced when the area under the receptive field is a “figure” compared to when it is “ground”. These authors attribute the enhancement to the presence of a figure border that would stimulate figure-ground segregation processes. This interpretation opens a new level of discussion because the identification of a region as a figure requires global image processing (the system needs to evaluate an area of the size of the figure or more), whereas feature contrast requires only processing of some neighborhood of the point in consideration, and even the findings concerning illusory contour representation might be explainable in terms of neighborhood processing (Heitger et al., 1998). Figure-ground segregation is fundamental to visual object recognition (Koffka, 1935), and finding this process reflected in signals at the level of V1 would require a new interpretation of visual cortical processing. However, two observations seem to limit the scope of this idea. One is the finding that the contextual enhancement decreased steeply with the size of the figure, reaching zero at ~8–10° figure size (Zipser et al., 1996). Figure-ground perception is not limited to small figures. The other limitation is the underlying assumption of a point-by-point representation of object surfaces (isomorphic coding). The figure-ground dimension is thought to be encoded by modulation of the activity evoked by texture elements. This may be plausible for textured objects, but not for objects of uniform color, because the vast majority of V1 neurons are not activated by uniform surfaces (Hubel and Wiesel, 1968; von der Heydt et al., 1996).

In the experiments to be described we have studied the context dependence of contrast border responses using displays in which the same local feature (contrast edge or line) was presented as part of a figure either on one or the other side of the receptive field. The Gestalt psychologists have pointed out that perception tends to “assign” contrast borders to objects (Koffka, 1935). Rubin’s famous vase figure demonstrates this compulsion of the visual system (Fig. 1A). The border is perceived either as the contour of a vase or as the contours of two faces. Figure 1B is generally perceived as a

Received March 13, 2000; revised June 15, 2000; accepted June 16, 2000.

This work was supported by National Eye Institute Grant EY02966. H.S.F. was supported by the Whitaker Foundation. We thank Ofelia Garalde, Hai Dong, and Clark Jefcoat for technical assistance and C. E. Connor, V. B. Mountcastle, E. Niebur, G. F. Poggio, and two anonymous reviewers for helpful comments on earlier versions of this manuscript. We acknowledge the use of an experimental setup and software developed by G. F. Poggio.

Correspondence should be addressed to Rüdiger von der Heydt, Krieger Mind/Brain Institute, Johns Hopkins University, 3400 North Charles Street, Baltimore, MD 21218. E-mail: von.der.heydt@jhu.edu.

Copyright © 2000 Society for Neuroscience 0270-6474/00/206594-18\$15.00/0

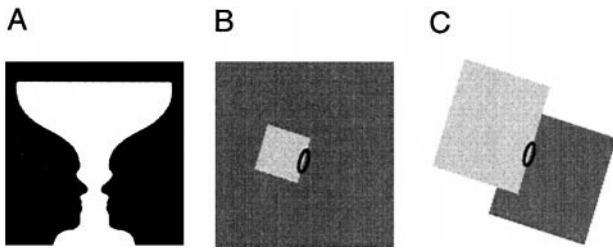


Figure 1. Perception of border ownership. *A*, Rubin's vase (Rubin, 1915). This well known ambiguous figure demonstrates the tendency of the visual system to interpret contrast borders as occluding contours and to assign them to one of the adjacent regions. In this example, figure-ground cues have been carefully balanced, but the black and white regions are generally not perceived as adjacent; instead, perception switches back and forth, and the borders belong either to the vase or to the faces. *B*, Isolated regions of contrast are generally perceived as "figures", that is, objects seen against a background. *C*, This display is generally perceived as two overlapping rectangles rather than a rectangle adjacent to an L-shaped object.

white square against a dark background (rather than a window in a dark screen) and the square "owns" the borders. When two regions are perceived as overlapping figures, the border between the two is owned by the overlaying figure (Fig. 1C). Our results indicate that this perceptual tendency to assign borders to objects is reflected in the neural activity at early cortical levels.

MATERIALS AND METHODS

Single neurons were recorded from areas V1, V2, and V4 in eight hemispheres of four alert, behaving monkeys (*Macaca mulatta*). The animals were prepared by attaching a peg for head fixation and two recording chambers (over the left and right visual cortex) to the skull with bone cement and surgical screws. The surgery was done under aseptic conditions under pentobarbital anesthesia induced with ketamine; buprenorphine was used for postoperative analgesia. All procedures conformed to the principles regarding the care and use of animals adopted by the American Physiological Society and the Society for Neuroscience, as verified by the Animal Care and Use Committee of the Johns Hopkins University.

Recording

The methods of recording were essentially the same as in von der Heydt and Peterhans (1989). Several weeks after the surgery, 1 or 2 d before the beginning of recording, a 3 mm trepanation was made in one of the chambers under ketamine anesthesia. On each experimental day, granulation tissue was removed from the dura, the hole was sealed with bone wax, and a microelectrode for extracellular recording was inserted through the wax and the dura mater, using a microdrive and positioning device mounted on the chamber. Electrode and wax were removed after the session, and dexamethasone drops were applied to reduce tissue reaction. Good recordings with minimal dimpling of cortex were usually possible for ~2 weeks after drilling a hole. After a break of ≥ 1 week, another hole was drilled, and recording resumed for up to five holes in each chamber. Electrodes with fine tips were used that easily isolate single cells (platinum-iridium, 0.1 mm diameter, taper 0.07–0.1 glass-coated, impedance 3–15 M Ω at 1 kHz; von der Heydt et al. 2000). While advancing the electrode, we monitored the entry into the cortex, the amount of single and multiunit activity, its orientation and ocular preference, the entry into the white matter, the entry into the cortex below the white matter, etc., and recorded their depths graphically. Comparison of many such track charts (~50 per hemisphere) with the histological reconstructions showed that layers 4B, 4C, and 6 in V1 can often be identified physiologically during the recording (von der Heydt and Peterhans, 1989).

Anatomical methods

After the recordings were completed, the animal was anesthetized, and thin, sharply pointed marker pins were inserted in parallel tracks at known positions around the recording regions with the positioning device used for recording. The animal was then given an overdose of pentobarbital, and the brain was perfused with buffered 4% formaldehyde. The pins were removed, the tissue was blocked and soaked in 30% sucrose, and 50 μ m frozen sections were cut at right angles to the orientation of the pins (tangential sections). The sections were stained for cytochrome oxidase. The positions of the recording tracks were determined from the electrode positioning coordinates by interpolating between the positions of the marker pins. For one animal (M12) the recording sites were reconstructed by tracing the outlines, layers, and pin holes with a computer-controlled microscope (NeuroLucida) and plotting the tracings together with the positions of the recording tracks. The depths were determined by aligning the depth records in the track charts (see above) with the corresponding anatomical landmarks. This method generally confirmed our previous

Table 1. CIE (1931) coordinates of the stimuli typically used for testing color selectivity

Color	x	y	Y (cd/m ²)
Red–brown	0.60	0.35	14–2.7
Green–olive	0.31	0.58	37–6.7
Blue–azure	0.16	0.08	6.8–1.8
Yellow–beige	0.41–0.46	0.50–0.45	37–6.5
Violet–purple	0.30	0.15	20–3.4
Aqua–cyan	0.23	0.31	38–7.3
White–gray–black	0.30	0.32	38–8.8–1.2
Light gray (background)	0.30	0.32	20 ^a –16 ^b

^aFor setup 1.

^bFor setup 2.

Y is the luminance, and x and y are the chromaticity coordinates. Two luminance levels were used for each chromaticity, except for the neutral colors, which had three luminance levels, and yellow and beige for which slightly different chromaticities were used.

identification of cortical layers according to physiological criteria. In the other three animals, only the locations of the tracks were determined to verify the cortical areas. In this case, the layer assignment of V1 was based only on the track charts.

Visual stimulation and behavioral paradigm

Two experimental setups were used, setup 1 for animals M12 and M15 and setup 2 for animals M13 and M16 (in the results to be presented below, the first digits of the neuron identification numbers indicate the animal). In setup 1, visual stimuli were generated by an Omnicomp GDS 2000 processor controlled by a personal computer and displayed on a Hitachi HM4119 color monitor with a 60 Hz refresh rate. Fixation target and test stimuli were viewed through a mirror stereoscope at a distance of 51 cm. The visual field measured 11.5° square for each eye, with a resolution of 400 \times 400 pixels. In setup 2, visual stimuli were generated by a Silicon Graphics Indigo2 workstation and displayed on a BARCO CCID 121 FS color monitor with a resolution of 1280 \times 1024 pixels and 72 Hz refresh rate. This display was viewed directly with both eyes at a distance of 93 cm and subtended 21° by 17° visual angle. The stimuli were colored or gray rectangles presented on a neutral gray background, as specified in Table 1. Eye movements were monitored by means of video-based infrared pupil tracking (Iscan) with 0.15° horizontal and 0.28° vertical resolution.

The animals were trained to fixate their gaze by requiring them to respond to an orientation change that could only be resolved in foveal vision. The fixation target was a 7 arc min white square divided by a thin gray line whose change from vertical to horizontal had to be detected. The target was centered on a 19 arc min black square to facilitate fixation. The general trial sequence was as follows: target onset, monkey responds by pulling a lever and begins to fixate, 0.5–5 sec variable interval (fixation period), target rotates, monkey responds by releasing lever, 1–2 sec variable interval (monkey usually looks away from target), new trial begins with target onset, etc. The hit rate during recording sessions was ~95% on average. Two of the monkeys (M13 and M16) served also in a study on perceptual filling in, in which they were trained to respond to a color change of a peripherally viewed disk-ring stimulus. There was no difference between these monkeys and the other two monkeys in the results of the present experiments.

Procedure. To study a representative sample of cells, an exhaustive analysis was attempted. We did our best to study every cell that was isolated and not to skip "difficult cells." After isolation of a cell, the receptive field was examined with rectangular bars, and the optimal stimulus parameters were determined by varying the length, width, color, orientation, and binocular disparity (in setup 1). Using this optimal stimulus, we then determined the "minimum response field", which was defined as the minimum visual field region outside which the stimulus did not evoke a response (Barlow et al., 1967). In other words, the bar has to enter this region to evoke a response. The size of the minimum response field characterizes the precision of positional information in the neural responses (see Results). This field is generally smaller than the area of summation that is apparent when stimuli of various sizes are tested. In cells that require a certain length of contrast border in the receptive field to respond, the "length" of the minimum response field can be negative (Henry et al., 1978). We have verified the accuracy of our maps by recording the position-response profiles parallel and orthogonal to the optimal orientation (see Figs. 11–13 for examples). Edge selectivity was measured by calculating the surface-to-edge response ratio for a square (usually 4°), defined as $(R_{\text{inside}} - R_{\text{outside}})/(R_{\text{edge}} - R_{\text{outside}})$, where R_{inside} is the response to the center of the square, R_{outside} the response outside the figure, and R_{edge} the maximum of the responses to the two optimally oriented edges of the square. (Note that a zero surface-to-edge response ratio does not necessarily mean zero surface response, but only that the responses for the inside and outside conditions

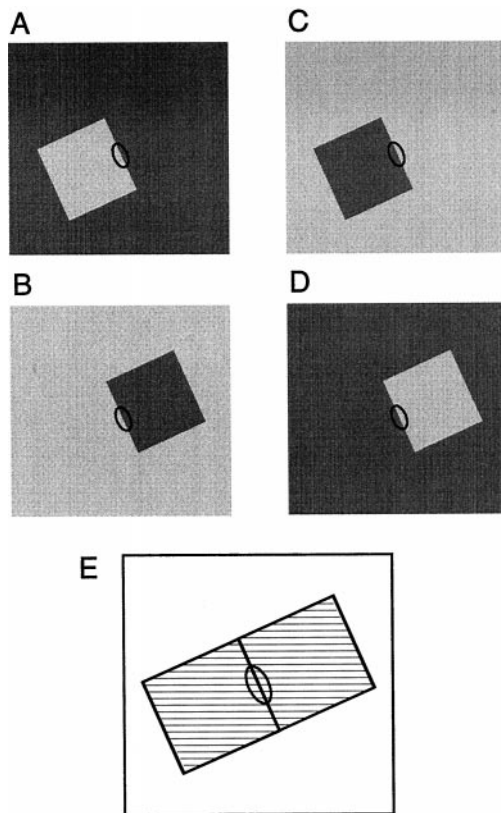


Figure 2. The standard test for determining the effect of border ownership on edge responses. In *A* and *B*, identical contrast edges are presented in the receptive field (*ellipses*), but in *A*, the edge is the right side of a dark square, in *B*, it is the left side of a light square. The relation is analogous between *C* and *D*, with reversed contrasts. *E*, The *hatched region* indicates the neighborhood of the receptive field in which displays *A* and *B* (or *C* and *D*) are identical. The preferred color of the cell (including *black*, *white*, and *gray*) and a *light gray* were used as the colors in these displays.

were equal. In most cells R_{outside} was zero or very low, as was the spontaneous firing rate.)

Standard test. Figure 2 illustrates the test used for determining the influence of border ownership^a on neural edge responses. A uniformly colored square was presented on a uniform background of a different color (we use the term “color” to include black, white, and grays). An edge of the square, at optimal orientation, was centered in the receptive field (represented by the ellipses in Fig. 2). Two colors were used, the previously determined optimal color (shown as dark gray in Fig. 2) and light gray. The optimal color was selected from a set of 15 colors (Table 1). As the Table shows, there was generally a luminance difference between the two colors. The colors of square and background, and the side of the square, were switched between trials, resulting in the four conditions shown in Figure 2. Note that the contrast borders presented in the receptive field in *A* and *B* are identical, but in *A* the border is the right side of a light square, and in *B* it is the left side of a dark square. This is similar for pair *C* and *D*, with reversed contrast. The neighborhood around the response field in which displays *A* and *B* (or *C* and *D*) are identical is defined by the size of the square, as illustrated by hatching in Figure 2*E*. In the standard test, sizes of 4 or 6° were used for cells of V1 and V2, and sizes between 4 and 17° were used for cells of V4, depending on response field size. In many cells, a range of sizes was tested. The four stimuli of Figure 2 were presented in counterbalanced sequences, for example *A-D-B-C-C-B-D-A*, to maintain color adaptation uniform and stationary and to control for possible changes in responsiveness. Initially, we have used static displays, changing the display between trials (when the monkey was not fixating). Thus, the square appeared before fixation and remained on throughout the trial. To study the time course of the responses, we have also used switching displays. In this case, a uniform screen of the color midway between figure and background colors was displayed during the intertrial intervals, and both figure and background were then turned on simultaneously ~300 msec after key pulling. This display remained on during the fixation period

^aIn images of three-dimensional scenes, the optical projection defines the borders between foreground and background regions; the borders are the occluding contours. Using the term “border ownership” in connection with two-dimensional displays, we refer to the relationship perceived by human observers. With the relatively simple displays used in this study there is generally no ambiguity.

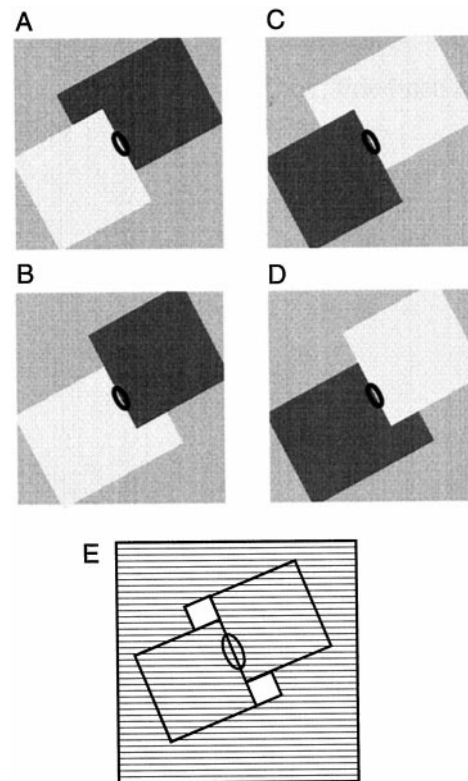


Figure 3. Overlapping figure test. In each of these displays two regions of approximately the same area are presented on either side of the receptive field (*ellipses*). As in Figure 2, the contrast edges in the receptive field were identical in *A* and *B* and in *C* and *D* but belonged perceptually to different figures. In *E*, the *hatched area* indicates the region of identical stimulation.

and switched back to the intermediate blank field after the monkey released the key at the end of the trial. By applying color changes to the figure and the surrounding area, this method of switching preserves the symmetry of the edge in the receptive field.

Overlapping figures. In this test, the border between a square and an L-shaped region was centered on the receptive field as illustrated in Figure 3. Human observers generally perceive such displays as two overlapping figures. Again, the contrast borders in the receptive field are locally identical in *A* and *B*, but belong perceptually to different figures. The same is true for *C* and *D*, with figure colors reversed. Most of the display remained unchanged between *A* and *B* (or *C* and *D*), as indicated by hatching in Figure 3*E*.

Data collection and data analysis

The signal from the microelectrode was passed through adjustable 24 db/octave high-pass and low-pass filters and a window amplitude discriminator. Spike events were recorded with 0.1 msec resolution, and those during the fixation period were analyzed. To determine the time course of responses, the spike trains of each cell were convolved with a Gaussian, averaged over repetitions, and normalized to the mean firing rate of the cell during the period of analysis (1 sec). For the curves of Figure 20 the normalized responses were averaged across cells for each cortical area (convolution with $\sigma = 16$ msec for V1, and $\sigma = 8$ msec for V2 and V4). To quantify the latencies, the point at half height between the level at stimulus onset and the peak of the convolved signal was determined ($\sigma = 8$ msec). Significance tests and analysis of the reliability of neural coding were based on mean firing rates during successive 1 sec intervals, beginning 300 msec after key pulling (or the time of figure onset in the case of switching displays). Significance of effects of border ownership and local contrast polarity was determined by ANOVA, and reliability of single-cell responses was assessed by determining the proportion of correct responses of a simple decision model, as explained in Results.

RESULTS

Because the representation of contrast borders was the focus of this study, only results from edge-selective neurons are reported in this paper. This means that all neurons included in this study (1) responded to lines or edges much longer than the receptive field (neurons with strong end stopping were excluded), and (2) did not respond, or responded much less, when a large uniform stimulus

was centered on the receptive field. For neurons of V1 and V2, edge selectivity was determined from the responses to a 4° (occasionally 6°) square, and the criterion was a surface-to-edge response ratio <0.25 (see Materials and Methods), which means that the cell responded at least four times more vigorously to an optimally oriented edge than to the center of the square. In V2 and the top layers of V1, ~85% of the neurons are edge-selective by this criterion (H. Zhou, H. S. Friedman, and R. von der Heydt, unpublished observations). In V4, edge selectivity was assessed with larger figures because the minimum response fields of most V4 cells were much larger than those of V1 and V2 cells (see below). For many V4 cells the minimum response field could not be determined because part of its boundary was outside the display field. Therefore, for V4 cells, the most sensitive position (“hot spot”), as determined with a bar or edge, was usually taken as the center of receptive field for the subsequent tests. Only cells with clear edge selectivity were studied. These were <50% of the V4 cells that we attempted to study. Our sample also includes a few cells that responded to thin bars, but to neither surface nor borders of uniform color squares. These cells were studied with outlined figures. The vast majority of the cells included in this study were orientation-selective (see below).

Border-ownership coding was studied in 206 cells, 63 of V1, 91 of V2, 45 of V4, and seven from the V1–V2 and V2–V3 border regions. The cells of V1 were recorded in layers 2 and 3 of the cortex. Fifty-six cells (30 of V1 and 26 of V2) were studied using the small-field, stereoscopic setup, and 143 cells (33 of V1, 65 of V2, and 45 of V4) were studied using the large-field, direct-view setup. Thirty-seven cells (20 of V1, 17 of V2) were studied with static displays, the rest with switching displays (see Materials and Methods). Solid squares as shown in Figure 2 were used if a cell responded to edges, which was generally the case (187 cells). If a cell responded only to lines and thin bars, outlined squares were used (19 cells; 2 of V1, 6 of V2, 11 of V4). (A number of cells were tested with both solid and outlined squares and with other types of figure-ground displays.)

The receptive fields of the cells of V1 and V2 were located in the lower contralateral visual field at eccentricities between 0.6 and 6° (median, 1.5°) for V1 and between 0.2 and 7.4° (median, 2.0°) for V2. The standard test was performed with squares of 4°, or sometimes 6° size. This is much larger than the typical size of the minimum response fields of V1 and V2 at those eccentricities. In V1, the length of the minimum response fields varied between 0.2° and 1.2° (median, 0.5°), and the width varied between 0.1° and 1.2° (median, 0.5°). In V2, the lengths were 0.2–3.0° (median, 0.7°), and the widths were 0.1–2.7° (median, 0.4°). For 13 cells (2 of V1 and 11 of V2) the “length” of the minimum response fields was negative (see Materials and Methods). Most of the cells tested in V1 and V2 responded best to long edges or bars; only 15 of 161 cells exhibited moderate degrees of end stopping (responses to short bars 2.1 times stronger than responses to long bars on average). In V4, the eccentricities of receptive fields ranged from 0.3 to 11° (median, 6.6°). In 10 of the V4 cells the standard test was performed with square sizes of 4–8° and in 24 cells with square sizes of 10–17°. One edge of the square was centered on the hot spot, while the other edge was either outside the response field or cutoff by the display. The lengths of the minimum response fields ranged from 1.7 to 12° (median, 3.6°), and the widths ranged from 0.1 to 5° (median, 2.4°).

In the following Section 1, we will present the results on the frequency and strength of border-ownership signals and contrast polarity signals, as assessed with the standard test, in the three cortical areas. We will discuss the range of spatial integration and the time course of the border-ownership signals and try to relate these findings to the conventional receptive field properties. In Section 2, we will present results of experiments in which we varied the stimulus configurations to gain insight into the mechanisms of border-ownership coding.

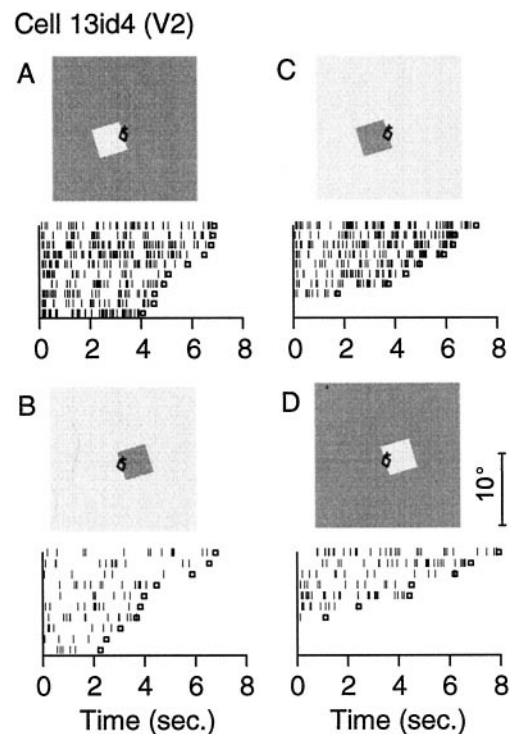


Figure 4. Example of border-ownership coding in a cell of area V2. The stimuli are shown at the top, and event plots of the corresponding responses are shown at the bottom. The ellipses indicate the location and orientation of the receptive field, and the crosses show the position of the fixation target. In the event plots, small vertical lines represent the times of action potentials, relative to the moment of lever pulling (which generally indicated the beginning of fixation). Small squares indicate the times of target flip (end of fixation). Each row represents a trial. Several repetitions are shown for each condition, sorted according to the length of the fixation period. *A, B*, The cell responded better to the edge of a green square on the left side than to the edge of gray square on the right side of the receptive field, although both stimuli were locally identical (green depicted here as light gray). *C, D*, When the colors were reversed, the cell again responded better to an edge that belonged to a square on the left than a square on the right. Square size, 4°; length of minimum response field, 0.4°; location in visual field (0.0°, –1.7°).

Section 1: results obtained with the standard test

Each neuron was tested with the four kinds of displays shown in Figure 2 (see Materials and Methods). If the activity of a neuron was determined by local features, it would respond equally to *A* and *B*, and equally to *C* and *D*, because these pairs of stimuli are locally identical. However, we found that many neurons responded to the same local edge differently, depending on the side to which the edge belonged. Based on the results of the standard test, we distinguished four types of results: (1) cells coding border ownership, (2) cells coding the polarity of edge contrast, (3) cells coding border ownership and polarity of edge contrast, and (4) cells coding neither border ownership nor polarity of contrast. We will first present examples of these four types of results and describe some control experiments, and then we will discuss the classification of cells and their reliability in signaling border ownership and contrast polarity.

Type 1: border ownership

Figure 4 shows the responses of a cell from V2. The stimulus was a green square surrounded by gray in *A* and *D*, and a gray square surrounded by green in *B* and *C* (the cell was not particularly color-selective, but green produced the largest response). The ellipses indicate the minimum response field of the cell, and the crosses mark the position of the fixation target. The raster plots at the bottom show the responses to repeated random presentations of the four stimuli. (Each row of small lines represents the activity during one fixation period; for each condition, responses have been sorted by the length of the fixation period.) It can be seen that the

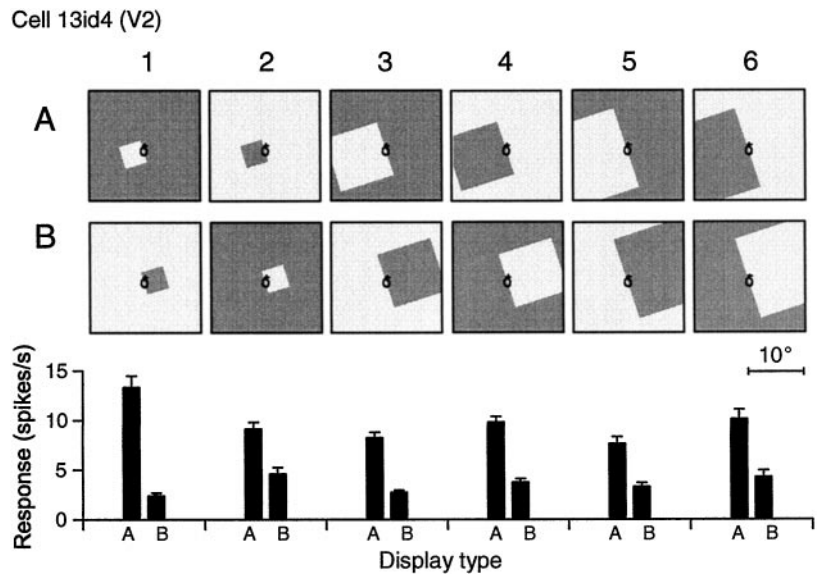


Figure 5. Size invariance of border-ownership coding. The same V2 cell as in Figure 4. Rows *A* and *B* show the stimuli, with pairs of locally identical stimuli juxtaposed. Conventions as in Figure 4. Bar graphs below show mean firing rates and SEs of the corresponding responses. Square sizes: 1 and 2, 4°; 3 and 4, 10°; 5 and 6, 15°. For each size, and for either contrast polarity, the responses were stronger when the square was located on the left side of the receptive field.

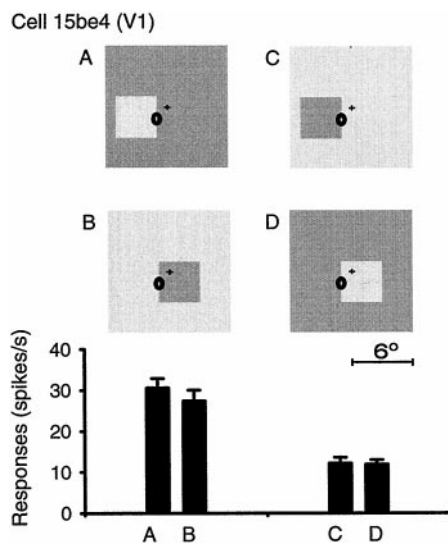


Figure 6. Selectivity for local contrast polarity. Cell of layer 2/3 of V1. The cell responded more strongly to light–dark edges (*A*, *B*) than to dark–light edges (*C*, *D*) irrespective of the position of the figure. The colors were yellow (depicted as light gray) and gray. Location of receptive field (–1.1°, –1.1°).

cell preferred a square on the left side (*A* and *C*) regardless of the figure contrast. Figure 5 shows the mean strengths of responses with SEs for the conditions of Figure 4 (4° squares) and for square sizes of 10 and 15°. Displays that were identical in the response field of the cell have been juxtaposed in rows *A* and *B*, and the bars below represent the corresponding responses. In each case, the cell responded more strongly when the square was on the left side (row *A*), and this preference was independent of the contrast polarity. Cells with this type of behavior were found in all three cortical areas, but more often in V2 and V4 than in V1.

Type 2: edge-contrast polarity

Cells that were selective for edge-contrast polarity were also observed in all three cortical areas. Figure 6 shows an example of a cell recorded in layer 2/3 of V1. Yellow squares on light gray background (*A*, *D*) were compared with light gray squares on yellow background (*B*, *C*). The cell responded more strongly in *A* and *B* than in *C* and *D*, but whether the square was on the left or the right side made no difference. Thus, the local edge contrast determined the responses of this cell.

Type 3: border ownership and edge-contrast polarity

Figure 7 shows the responses of a cell of V2 that was selective for side of figure and local contrast polarity. This cell was color-selective, preferring dark reddish colors, as illustrated in Figure 8. Brown and gray were used for this test, as shown. The cell responded to the top edge, but not the bottom edge of the brown square and barely at all to the edges of the gray square. The differences between *C* and *A* and between *D* and *B* indicate that the cell was selective for edge-contrast polarity, which is typical for many V1 simple cells (Schiller et al., 1976). However, responses were much stronger in *C* than in *D*, in which the local edge contrast

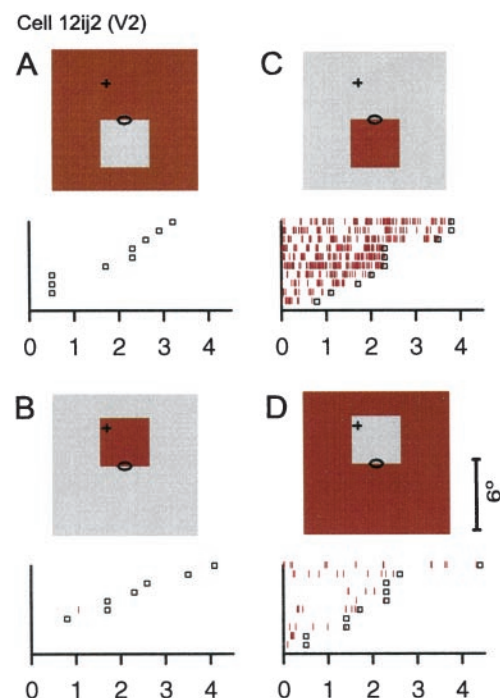


Figure 7. Example of simultaneous coding of border-ownership and edge-contrast polarity. This cell of area V2 was color-selective with a preference for dark, reddish colors (see Fig. 8). Brown and gray were used for the test. Conventions are the same as for Figure 4. The cell responded to the top edge of a brown square (*C*), but hardly at all to the bottom edge of a gray square (*D*), although in both cases the same gray–brown color boundary was presented in the receptive field. The cell did not respond at all to edges of the reversed contrast (*A*, *B*). Square size, 4°; length of minimum response field, 1.4°; location in visual field (1.4°, –3.0°).

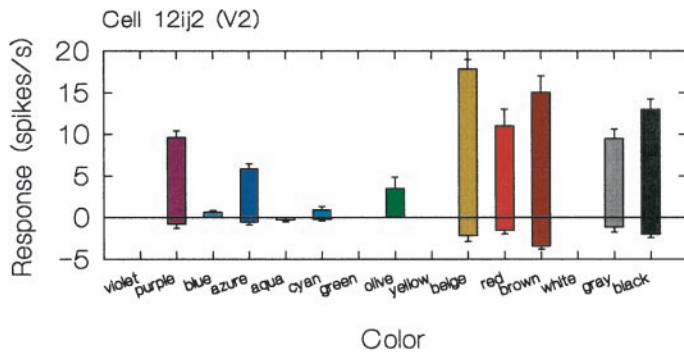


Figure 8. The color selectivity of the cell of Figure 7. Bars of 15 colors (Table 1) were flashed in the receptive field for 500 msec with intervals of 500 msec. The graph represents mean firing rates with SEs. Activity during the On phases is plotted upward, and activity during the Off phases is plotted downward. It can be seen that the cell responded to reddish hues better than to greenish hues, with a preference for the darker representative of each hue (*brown > red, beige > yellow, black > white*, etc.). Whereas, in most cells, the demonstration of border-ownership coding did not require chromatic stimuli, cells with striking chromatic selectivity were also common.

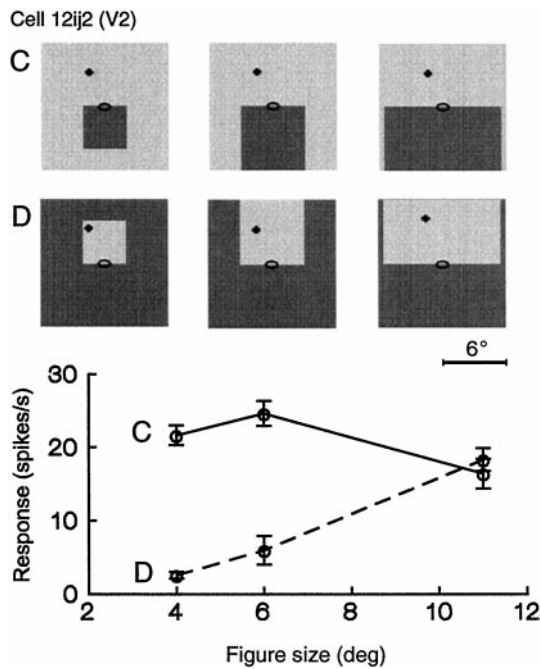


Figure 9. Size invariance of border-ownership coding in the color-selective cell of Figures 7 and 8. Side of ownership produced response differences for 4 and 6° squares, but not for squares of 11° size. Human observers also find the distinction of figure and ground weak for the largest size. Colors were *brown* and *gray*, here depicted as *dark* and *light gray*.

was the same, showing that the cell was influenced by a region much larger than the response field. Figure 9 shows that similar results were obtained with 4 and 6° squares. With 11° squares, which were partially cut off by the screen, the cell responded equally well to both stimuli. The reader may also find the figure-ground distinction weak for these displays. For configurations *A* and *B*, no responses were obtained at any size of the square. Cells with combined selectivity for border ownership and local contrast selectivity were found in all three cortical areas.

Type 0: neither border ownership nor contrast polarity

Many cells responded approximately equally to all four conditions of Figure 2. These were edge-selective and generally orientation-selective cells with no preference for local contrast polarity or figure-ground direction.

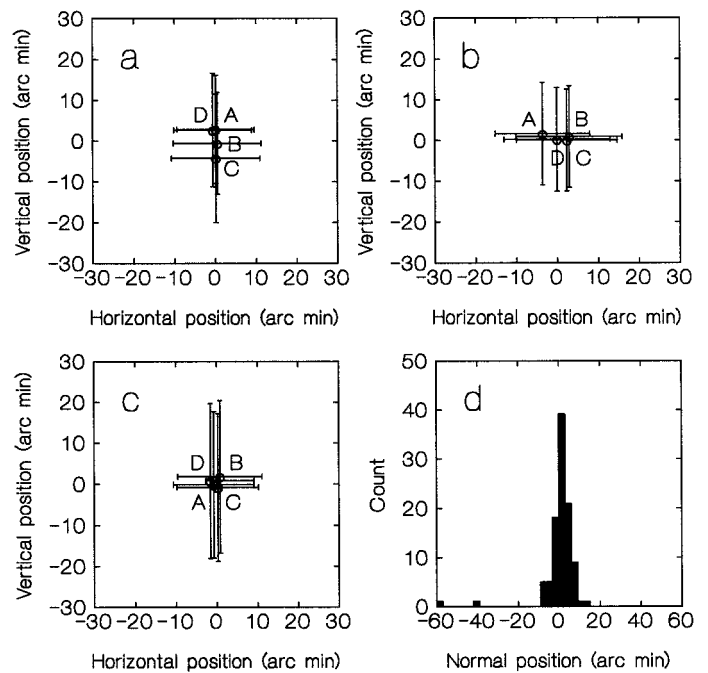


Figure 10. Eye movements during fixation. Recordings from monkey M16 during the standard test displays with 4° squares. *a–c*, Means and SDs of horizontal and vertical positions of gaze, grouped according to display type (Fig. 2, *A–D*). Each plot represents ~320 trials with stimulus orientations of 135° (*a*), 45° (*b*), and 135° (*c*) recorded in succession. If the side of the square had influenced gaze position systematically, means *A* and *C* would appear displaced relative to means *B* and *D* in the direction perpendicular to stimulus orientation. No systematic displacements were apparent. *d*, Histogram shows the distribution of the effects of figure position on eye movements perpendicular to stimulus orientation, as determined by ANOVA, in 101 blocks of 40 trials each. Positive values designate eye movements that would have moved the receptive field toward the center of the figure. Binwidth, 3 arc min. The eye movements were small and not systematically related to the side of the figure.

Controls

Given the generally small size of the minimum response fields of the cells, we wondered whether variations in retinal stimulus position caused by fixational eye movements could have affected the responses. Random variations would only add noise to the data, but if changes in fixation were related to the figure position, a systematic variation of responses could result. We have analyzed the eye movements quantitatively in one of our monkeys, M16. The eye movement displays and performance in the fixation task of the other animals indicated that their fixation behavior was similar. The following data were obtained during the 3 d after the neural data collection was completed. The same fixation task was used, and the four displays of the standard test were presented in counter-balanced order, exactly as during the neural data collection. Eye movements were recorded for ~4000 fixation periods, each 0.5–5 sec in duration. As always, only trials in which the animal performed the fixation task correctly were analyzed. Figure 10 shows the results. Plots *a–c* show data from three successive blocks of recording, each consisting of ~320 trials, with figure orientations of 45, 135, and 45°, respectively. Each plot represents the means and SDs of the eye position signal for the four stimulus conditions labeled *A–D*, as in Figure 2. It can be seen that the differences between the means were small compared to the overall variation. The relevant measure is of course the variation of gaze position perpendicular to the edge of the figure, that is, along the 135° diagonal for plots *a* and *c* and along the 45° diagonal for plot *d*. Thus, the question is whether means *A* and *C* were separated from means *B* and *D* along these axes. No such separation was apparent in these samples. To determine the statistical distribution of gaze shifts that might have influenced our analysis of neural responses, we have segmented the eye movement recordings into blocks of 40

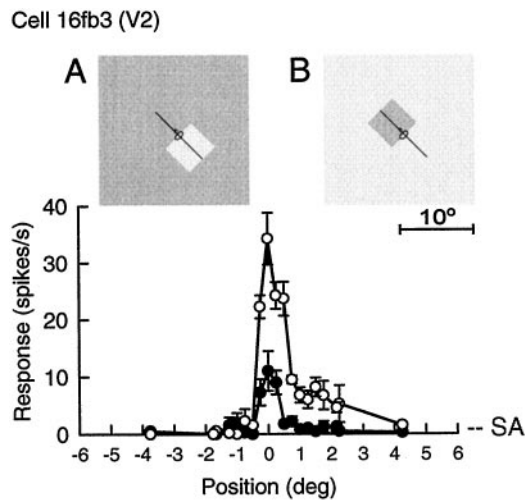


Figure 11. Position invariance of border ownership coding in a cell of V2. The top left edge of a light square (*A*), and the bottom right edge of a dark square (*B*), were centered on the receptive field, and position of the squares was then varied. Mean firing rates and SEs are plotted as a function of edge location relative to receptive field center. Open circles represent responses to edge of light square (*A*), and filled circles represent responses to edge of dark square (*B*). In either case, the maximum response was obtained when the edge was centered on the receptive field, but the responses were stronger for *A* than for *B* at any position. SA, Level of spontaneous activity. Ellipse, Minimum response field. Cross, Fixation target. Line indicates range of variation of edge position, positive toward bottom right. Colors were aqua and gray; size of square, 4°; receptive field location (0.4°, -1.7°).

trials, which was the typical length of blocks used to assess border-ownership selectivity. The component of eye movements in the critical direction was computed, and a two-way ANOVA with the factors side-of-figure and contrast polarity was performed for each block. The histogram in Figure 10*d* shows the distribution of the effects of the factor side-of-figure. One can see that the gaze shifts were generally small (absolute size <2.8 arc min in 50%, <7 arc min in 95% of the cases) compared to the size of the minimum response fields of the cells that we have studied (median widths of 30, 24, and 144 arc min for V1, V2, and V4). Thus, the effects of variations in gaze position on the magnitude of the neural responses were probably insignificant.

Because the retinal image is never stationary under normal conditions, border ownership signals should not depend critically on stimulus position if they were to serve a function in vision. Figures 11 and 12 show examples of recordings in which we have studied the effect of varying the stimulus position. The cell of Figure 11 was recorded in V2 and had a small response field. The cell of Figure 12 was from V4 and had a large response field. For each cell, an edge was positioned in the receptive field center at the preferred orientation and presented either as part of a dark square or as part of a light square, as shown in the insets. Position was then varied along an axis orthogonal to the edge. The responses are plotted as a function of edge location relative to the center position. Open circles represent the responses obtained with the light square (*A*), filled circles represent the responses obtained with the dark square (*B*). It can be seen that the cell of Figure 11 responded better to the top left edge of the light square than to the bottom right edge of the dark square at any position, and the cell of Figure 12 responded to the top edge of the light square, but not the bottom edge of the dark square, at any position. These results show that border ownership signals are robust and do not depend on the exact positioning.

As pointed out above, the size of the square determines the area of identical stimulation in the standard test. In other words, a difference in response indicates an influence from outside the region occupied by the squares in the two positions (Fig. 2, *hatched area*). This area was generally much larger than the minimum response field of the cell. It is important to note that in cells of V1

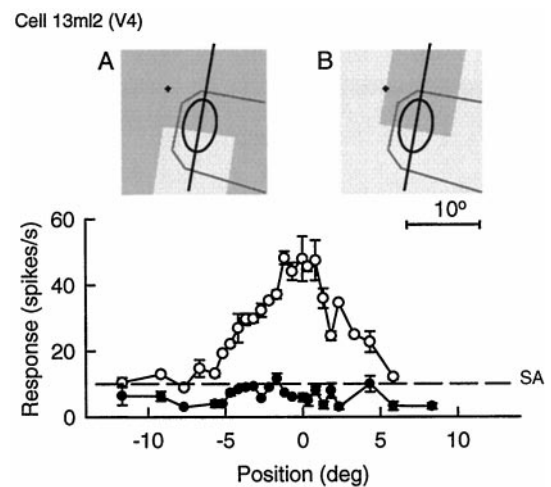


Figure 12. Position invariance of border ownership coding in a cell of V4. The cell responded to the top edge of the light figure (*A*), but not the bottom edge of the dark figure (*B*) at any position. SA, Spontaneous activity. Straight line indicates range of variation of edge position, positive downward. Ellipse delineates the most sensitive region of the receptive field (approximate contour of half-maximal response), the gray line the more uncertain total extent (note that the short axis of the ellipse corresponds to the preferred orientation of the cell). Colors, yellow and light gray; width of figure, 10°; receptive field location (4.6°, -4.9°).

and V2 and also in many cells of V4, the minimum response field is sharply defined, and no response can be evoked with bars outside this field. Figure 13 illustrates an example of a cell of V2 with a small receptive field near the fovea. *A* shows position response functions obtained with an optimally oriented bar of 1° length. The insets illustrate bars and receptive field. It can be seen that the responses drop to zero when the bar leaves the small region marked by the ellipse. This region measures <1° along the preferred orientation. The limits of the response field were confirmed with a square of 8° size (*B*). In the standard test, performed with squares of the same size, the cell responded better when the figures were on the bottom left side of the receptive field than on the other side (*C*). Thus, although contours presented as close as 1° to the receptive field center produced zero responses, the activity evoked by a contour in the response field was strongly modulated by the image context at a distance of >4°. Note also that the cell was not end-stopped (strong response to the center of the 8° long edge).

Quantitative classification

We found no differences regarding the coding of border ownership between the data recorded with the two setups. Specifically, the proportions of cells that showed an influence of side-of-ownership were similar. Also, tests with displays that remained constant throughout the trial and displays in which the figures were switched on after the beginning of fixation produced essentially the same results (Fig. 14). The data from the two setups and the results obtained with static and switching displays were therefore pooled for the statistical analysis.

Figure 15 shows, for the three cortical areas, the distributions of the magnitude of the effect of side-of-ownership, expressed as the ratio between the responses to the nonpreferred side and the preferred side. For cells tested with the standard solid squares ($n = 180$) the responses for the two contrast polarities were added, for example, $(B + D)/(A + C)$ for the cell of Figure 4. Another 19 cells that were tested with outlined squares (in which figure and background had the same color) are included in this figure. It can be seen that response ratios <0.5 were common in V2 and V4. For cell 13id4 of Figure 4 the ratio was 0.35, for cell 12ij2 of Figure 7 it was 0.11. Cells with such low ratios were also found in V1, but less frequently. The statistical significance of the effects of border ownership and contrast polarity was determined by performing a three-factor ANOVA on the data of the standard test for each cell,

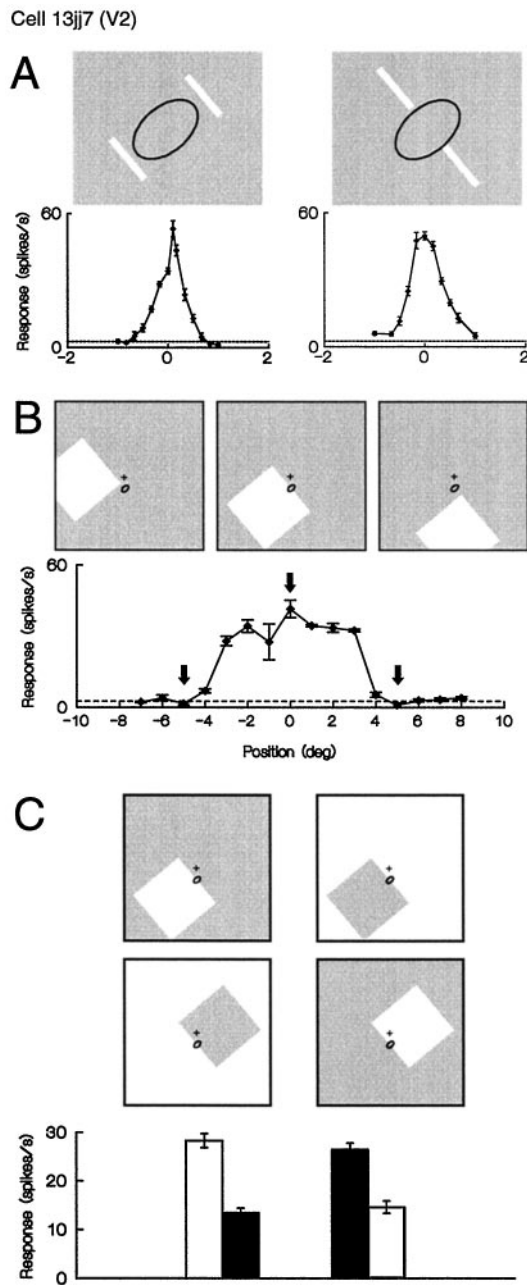


Figure 13. Comparison of conventional receptive field size and extent of image context integration. Responses of a cell of V2. *A*, Firing rate as a function of the position of a 0.2° wide, 1° long static white bar on gray background. The bar was presented at the preferred orientation of the cell, and position was varied along that orientation (right) and along the orthogonal axis (left). *Insets* show the bar at positions +1 and -1°, corresponding to the end points of the plotted curves. *Ellipse* indicates the region outside which the bars did not evoke a response (“minimum response field”); note that the preferred orientation is that of the short axis of the ellipse). *Dashed lines* indicate level of spontaneous activity. *B*, Responses to an edge of a static square of 8° size at various positions along the preferred orientation. Stimulus displays are illustrated for three data points (arrows). The responses were approximately constant as long as the edge remained inside the minimum response field and dropped to zero when the edge left the field. *C*, Test for border ownership. *Open bars*, Responses to white squares; *filled bars*, responses to gray squares. Despite its small receptive field, the cell differentiated displays that were identical in an 8 × 16° region around the receptive field. Note different scales of stimulus insets in *A* versus *B* and *C*. Colors, white and light gray; receptive field location (0.2°, -1.7°).

the factors being side-of-ownership (Fig. 2, *A* and *C* vs *B* and *D*), local contrast polarity (*A* and *B* vs *C* and *D*), and time after lever pulling (which indicates the beginning of fixation). Only the spikes during the fixation period were included, and data were sampled in

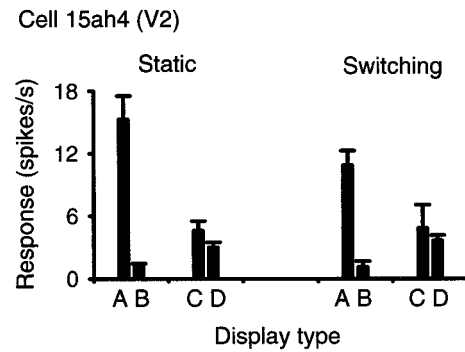


Figure 14. Static and switching displays produced similar results. Responses from a cell of V2 that was selective for border ownership and contrast polarity. The labels refer to the displays of Figure 2. Colors red and gray were used.

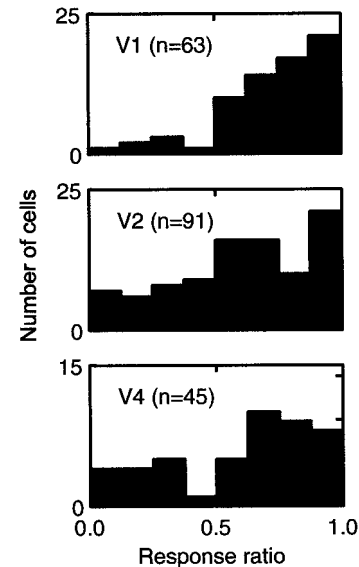


Figure 15. The distributions of the magnitude of the border-ownership effect in the three cortical areas V1, V2, and V4. The response ratio is the ratio of the mean response to the nonpreferred side over the mean response to the preferred side.

1 sec bins. The factor time was included to account for response variations during the fixation periods. A significance level of 0.01 was chosen. The distributions of the four response types in the three cortical areas are summarized in Figure 16. In V2 and V4, more than half of the cells showed significant border ownership modulation (type 1 and type 3: V2, 59%; V4, 53%), compared to 18% of the cells of V1. A majority of these cells were also affected by the polarity of edge contrast (type 3). Cells that were influenced by border ownership irrespective of the contrast polarity (type 1) were rare in V1, but made up 15% of the cells in V2 and V4. The *p* values for side-of-ownership differed significantly between the areas ($p < 10^{-4}$; Kruskal-Wallis); the differences V1-V2 and V1-V4 were significant ($p < 0.005$), but the difference V2-V4 was not ($p > 0.75$). There was no difference between areas in the *p* values for local contrast polarity ($p > 0.17$).

Only the main effects of side-of-ownership and local contrast polarity were represented Figure 16. In many cells (58 of 180) the analysis indicated significant interaction of these factors. Most of these cells (38 of 58) were of type 3 (significant effects of side-of-ownership and local contrast polarity). To understand this result, consider the four displays of Figure 2. An absence of interaction implies that the sum of the responses to *A* and *D* equals the sum of the responses to *B* and *C*. For the cell of Figure 4, this was approximately true (interaction -9% of grand mean, NS), but not

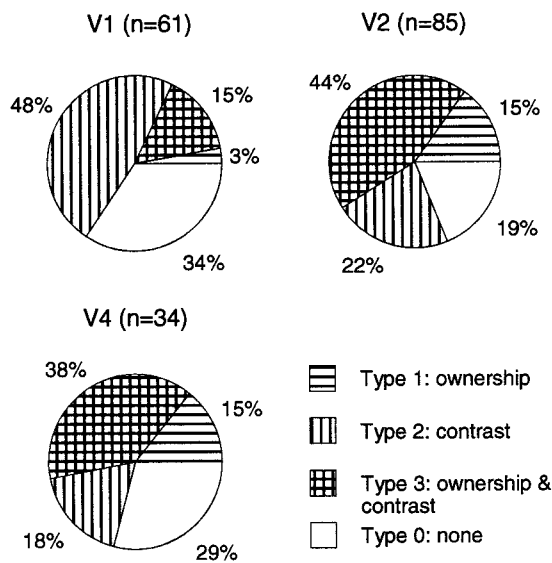


Figure 16. The distributions of the types of contour responses found in cortical areas V1, V2, and V4. Classification based on two-factor ANOVA. *Ownership*, Responses modulated according to side of ownership; *contrast*, responses modulated according to local contrast polarity; *ownership & contrast*, modulation by either factor; *none*, no modulation. In V2 and V4, more than half of the cells showed border-ownership modulation. Note that there are fewer cells in this figure than in Figure 15 because cells tested only with outlined figures are not included here.

for the cell of Figure 7, which responded almost exclusively to display C, so that the sum of B and C was much greater than the sum of A and D (interaction, +80%; $p < 0.001$). When type 3 cells showed interaction, it was nearly always positive, as in the example of Figure 7 (34 of 38; interaction, +17 to +96%; median, +46%). Some of these cells combine border ownership and local contrast polarity in the manner of an AND gate, responding exclusively to the edges of a figure of the preferred color, located on the preferred side.

The cell of Figure 7 responded to the edge of a brown figure embedded in gray (C) but not to the edge of a gray figure embedded in brown (D). Brown was the preferred color of the cell, as determined with flashing bars (Fig. 8). The question is of interest whether type 3 cells in general preferred displays in which the preferred color was figure over displays in which the preferred color was ground. This was in fact the case in 31 of 34 color or luminance-selective type 3 cells, while three cells preferred the displays in which the preferred color was ground.

Significant interaction between side of ownership and local contrast polarity was also found in 20 cells in which only one of the main effects, or none of them was significant (side of ownership, 6; local contrast polarity, 8; none, 6). This behavior is also interesting because these cells were edge-selective, but preferred figures of one color, no matter on which side the figure was presented. In the extreme, such a cell would respond, for example, to A and D of Figure 2, but not to B and C, although exactly the same edges were presented. What mattered was which color was figure and which was ground. Five cells of this kind were recorded in V1, eight in V2, and seven in V4.

The time after beginning of fixation had a significant effect only in a minority of cells (V1, 11 of 61; V2, 12 of 85; V4, 11 of 34). Thus, the mean firing rates of most cells were rather constant throughout the fixation period. Furthermore, there was generally no interaction between side-of-ownership and time (significant interactions were found in one cell of V1, two cells of V2, and three cells of V4). This means that, generally, the border-ownership signal neither strengthened nor weakened during the fixation period. Indeed, a two-factor analysis of the effects of border-ownership and contrast polarity, ignoring the time variable, produced similar results as the ones shown in Figure 16. Also the

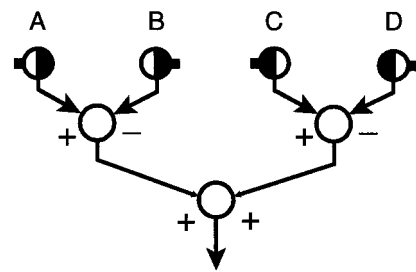


Figure 17. A decision model used to estimate the reliability of border ownership signals. The model assumes four neurons A–D with identical receptive fields except for reversals of side preference and contrast polarity preference, as indicated by the symbols (tabs for side preference, fill pattern for contrast polarity preference). The decision was based on the responses of these four neurons combined in the form $(A - B) + (C - D)$. Because the four model neurons have otherwise identical response properties, recording the responses of one of them to the four stimuli of Figure 2 provides the responses of all four neurons to any one stimulus. The analogous model was used for local contrast discrimination; in this case the decision was based on the signal $(A - C) + (B - D)$. See Results for explanation of how reliability estimates were calculated.

three-way interaction, which would indicate a time dependence of figure-color preference, was rarely significant (one cell of V1, two cells of V2, three cells of V4; three of these cells were among the aforementioned group with interaction between side-of-ownership and time). This analysis was done at a coarse scale, the exact time course of the response onset will be discussed below.

Reliability of signals

The above classification gives the answer to the question: was there a significant influence of the factors side of ownership and contrast polarity? A somewhat different question is: how well would the neural signals discriminate side of ownership and/or contrast polarity? This is the question of how reliable the information is that a group of cells can provide to other stages of processing. The first question examines the strength of the given data (and the answers depend on the amount of data sampled from the individual cells), the second question concerns the reliability of the neural signals (and the answer should not depend on sample sizes, except for an error of estimate).

The reliability of the neural border ownership signals was assessed by determining the proportion of consistent responses of a simple decision model, as shown in Figure 17. We assume that for each cell A recorded in a given area there exist three sister cells B–D, whose receptive fields are identical except that those of A and D are mirror images of each other, and so are those of B and C and that local contrast selectivity is reversed between A and C and between B and D. Note that we do not make assumptions about the actual selectivity of these cells. We only assume that for each receptive field structure that is realized there is also the mirror-image structure, and for each pattern of contrast polarity-sensitive inputs (e.g., ON- and OFF-types) there is also a pattern of reversed polarity. Because the cells have otherwise identical properties, recording the responses of one cell to the four stimuli of Figure 2 tells us how each of the four cells would respond to these stimuli. Our model figure-ground mechanism combines the responses of the four cells in the form $(A - B) + (C - D)$. (Note that the order of subtraction and summation can be interchanged.) If this signal is >0 , it is decided that the figure is located to the left of the receptive field, if it is <0 , it is decided that the figure is located to the right. This is an opponent model; we assume that the relative strength of responses of pairs of cells codes border ownership. Opponent models have been used to explain perceptual data on discrimination of orientation (Regan and Beverley, 1985), direction of motion (Newsome et al., 1989), and other perceptual dimensions. Summing over contrast polarity has the virtue of eliminating a possible effect of figure color, but is not essential. The example of Figure 4 shows that summation over contrast polarity occurs. In this case, our model would be equivalent to a simple

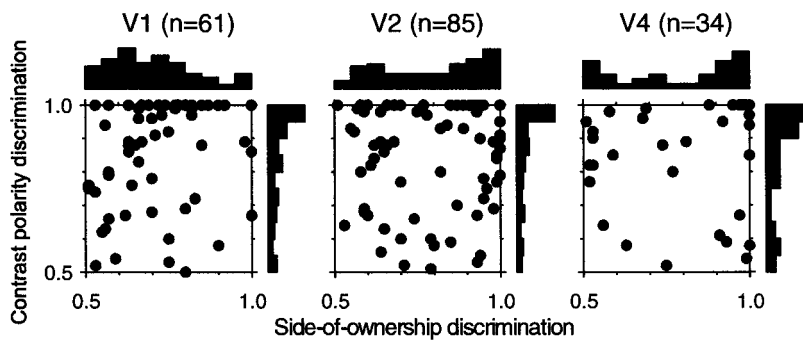


Figure 18. The reliability of neural responses in signaling border-ownership and local contrast polarity. Each *dot* represents the reliability estimates for one neuron derived from 1 sec samples of activity as explained in Results (0.5 = random, 1.0 = perfectly consistent). The histograms show distributions of reliability estimates for each dimension. Cells that reliably signaled edge-contrast polarity were common in V1 (*dots* at top of scatter plot). Their relative frequency was similar in the extrastriate areas, as shown by *histograms* on right margins. Cells that signaled border ownership were rare in V1, but common in V2 and V4 (*dots* near right margins). Substantial fractions of cells signaled both border ownership and polarity (*dots* in top right corners).

opponent model. We do not necessarily assume that single neurons are connected in the manner of Figure 17; the equivalent operation could be implemented with pools of cells. Our model is rather a tool for assessing the reliability of neural signals in ownership coding. The analogous model was used to assess the reliability of contrast polarity signals. It combines the responses in the form $(A - C) + (B - D)$, and a signal >0 is taken to mean light–dark edge, a signal <0 to mean dark–light edge.

Reliability estimates were obtained for each cell using the first second of responses. We calculated the output of the decision model for all combinations of the available responses to stimuli A–D of Figure 2 (on the order of 10^4 combinations). The proportion of cases in which the output indicated side of figure (or edge-contrast polarity) correctly was then determined (“correct” being defined by the mean responses to the four stimuli). This proportion estimates the probability that a correct decision would be made from 1 sec of response of the quadruple of cells to a random presentation of one of the stimuli of Figure 2 (assuming that the signals of the four cells are statistically independent). The resulting values are distributed between 0.5 (random) and 1 (perfectly consistent). To illustrate the results: the reliability of the cell of Figure 4 (type 1) was 97% for side of ownership and 82% for local contrast polarity; the cell of Figure 6 (type 2) was only 77% reliable for side of ownership, but 99% for local contrast polarity; the cell of Figure 7 (type 3) was 100% reliable for side of ownership, and 100% for local contrast polarity.

Figure 18 shows the joint distributions of the reliability estimates for border ownership and local contrast polarity discrimination for areas V1, V2, and V4. Each point represents an individual cell. Histograms of the integrated distributions are shown on the margins. The plots show that cells with perfect border-ownership coding (data points near the right margin) were common in V2 and V4, but rather rare in V1, where only four cells came close to 1. Nevertheless, the existence of such signals in the primary visual cortex is remarkable. The histograms at the top suggest that the distributions for V2 and V4 might be bimodal. Comparing medians, significant differences were found between V1 and V2 and between V1 and V4 ($p < 0.003$), but not between V2 and V4 ($p = 0.87$; Kruskal–Wallis). On the other hand, it can be seen that coding of edge-contrast polarity was excellent in many cells of V1 (data points near the top margin). Contrast polarity-selective cells were also found in the other two areas in similar proportions, as indicated by the histograms on the right margins. The median reliability in contrast polarity coding was not different between the three areas ($p > 0.5$). It can be seen also that some cells of V2 and V4 performed well in both dimensions (points in the top right corner).

Size invariance

The results of Figures 5 and 9 showed that neurons were border ownership-selective over a range of figure sizes. Because the size of the figure determined the area of the visual field that received identical stimulation in the displays to be compared (Fig. 2), the maximum size at which a reliable difference is still obtained indicates the extent of visual context integration and the extent of neural convergence which is necessary to achieve such integration.

We have examined 26 cells (three of V1, 17 of V2, and six of V4) with the standard test using different square sizes. Figure 19 plots the reliability of border-ownership coding as a function of size. Lines connect the data points of individual cells. The plots include results for displays in which large parts of the squares were cut off by the screen margins as, for example, in Figure 5. However, data for perceptually ambiguous displays have been excluded (that is, displays divided in two by a straight border, the displays on right of Fig. 9, and displays in which only one corner of the square was visible). Surprisingly, in most cases the reliability of border-ownership coding diminished only slightly with increasing figure size.

Time course

Response histograms were computed for the four conditions for each cell, normalized, and then averaged over cells (see Materials and Methods). Only cells with significant border-ownership modulation were included. The mean responses for the preferred and nonpreferred sides of the square (average of the two contrast conditions) are plotted in Figure 20 as functions of time after stimulus onset. It can be seen that the responses to the preferred and non-preferred sides diverged almost from the beginning in all three cortical areas. To quantify the latencies, we have determined the points of half-maximal signal for the sum and the difference of the responses for the two sides. The latencies of the summed

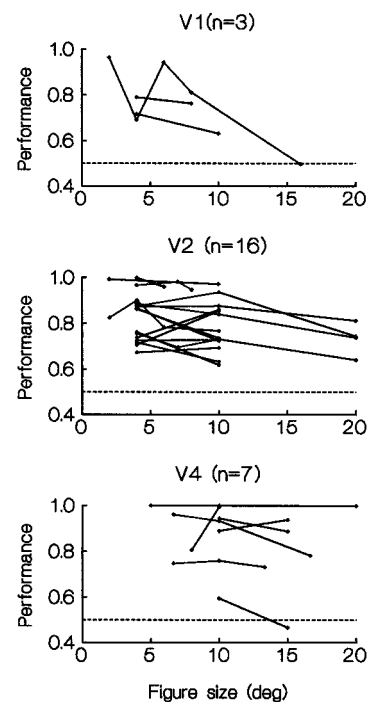


Figure 19. The effect of figure size on border ownership discrimination in cortical areas V1, V2, and V4. Each *point* represents a reliability estimate based on a test with one size; *points* for the same cell are connected by *lines*. The figure size generally had little effect.

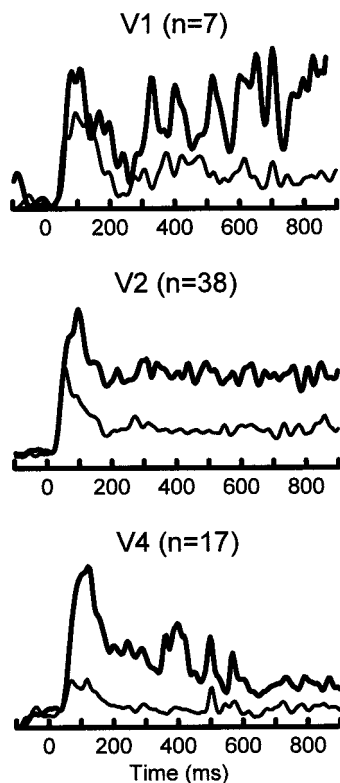


Figure 20. The time course of border-ownership modulation. The figure shows the average responses of all neurons with significant border-ownership modulation in the three areas. The responses of each cell were normalized to its mean firing rate during the fixation period and averaged. Zero on the time scale refers to the onset of the figure-ground display. *Thick* and *thin lines* represent responses to preferred and nonpreferred sides, averaged over both contrast polarities. A differentiation was evident shortly after the response onset for cells in all three cortical areas.

responses were 57 msec in V1, 43 msec in V2, and 63 msec in V4, and the corresponding latencies of the differences were 69, 68, and 73 msec. (The reason why the summed response of V2 has a shorter latency than that of V1 might be the higher relative weight of magnocellular input in the sample of V2 compared to that of top layer V1) (cf. Bullier and Nowak, 1995). Thus, a differentiation of side of ownership started within 10–25 msec after the onset of the responses. Because this differentiation depends on the processing of an image region that is at least as large as the figure and thus requires the transmission of information over some distance in the cortex, these short delays are surprising.

We wondered if other signals existed in the cortex that could provide the peripheral information earlier. Using the same method, we have analyzed the responses of the cells that did not differentiate side of ownership and obtained latencies of 54, 74, and 69 msec for the three cortical areas, which is similar to the latencies of the border-ownership cells. Taken together, these results indicate that the cortical processing that leads to border-ownership discrimination requires no more than ~25 msec.

Relation to conventional receptive field properties

Because the signals of some cells carry border-ownership information, whereas others do not, the question arises if this phenomenon can be related to any of the conventional receptive field properties. We have analyzed orientation, color, and disparity selectivity, and the property of end stopping. Orientation selectivity was quantified by the orientation modulation index:

$$\text{OMI} = \frac{R_{\max} - R_{\min}}{R_{\max} + R_{\min}},$$

where R_{\max} and R_{\min} are the maximum and minimum responses of the orientation tuning. As mentioned above, most cells included in this study were orientation-selective (OMI > 0.6; V1, 89%; V2, 78%; V4, 82%). The orientation modulation index was not different between cells with border-ownership selectivity (types 1 and 3) and other cells (types 2 and 0) ($p = 0.99$; Kruskal–Wallis).

Color selectivity was determined by using a set of 15 colors (Table 1) and computing a color selectivity index (CSI) according to:

$$\text{CSI} = 1 - \frac{\sum R_i}{nR_{\max}},$$

where R_i is the response to color i , R_{\max} the response to the preferred color, and n is the number of colors. For a cell that responded equally to all colors, the index would be 0, for a cell that responded only to one of the 15 colors, the index would be $1 - 1/15 = 0.93$. For cell 12ij2 of Figure 8, for example, we obtained a CSI = 0.67. Of 150 cells for which we have color selectivity data (52 of V1 and the V1/V2 border region, 69 of V2, and 28 of V4), 64 cells (43%) were color-selective (CSI > 0.5). There was no difference between cells with border-ownership selectivity and other cells ($p = 0.33$; Kruskal–Wallis). These results do not depend critically on the choice of the color selectivity index. A more extensive analysis showed that other criteria produced similar proportions of color-coded edge-selective cells (H. Zhou, H. S. Friedman, and R. von der Heydt, unpublished observations).

As mentioned above, cells with strong end stopping were excluded from this study because we were interested in the coding of figures with long, straight edges (for which ownership is ambiguous in the local perspective). Length response curves for moving bars were recorded from 80 cells of our sample. Length inhibition (long bar response less than half the optimum length response) was found in 8 of 36 (22%) of the cells with border-ownership coding, compared to 7 of 44 (16%) of the other cells, which is not significantly different. Thus, the presence of length inhibition is not an indicator of border-ownership coding.

The proportion of disparity-selective cells was about the same among border-ownership cells (5 of 10) and other cells (15 of 28). However, we do not have enough data for a detailed comparison of the types of disparity tuning. An interesting question is the relationship to stereoscopic edge selectivity, which will be discussed below.

In conclusion, the occurrence of border-ownership selectivity was not related to conventional selectivity for orientation, color, length of contour, or binocular disparity.

Section 2: results obtained with other displays and the question of figure-ground cues

Perceptual studies have shown that various factors determine whether a region is perceived as figure or ground. The presumed “cues” include figure convexity (convexity in the sense that a straight line connection of any two points of the figure is contained within the figure), closure of contour, occlusion features, and binocular disparity (Nakayama and Shimojo, 1990; Finkel and Sajda, 1992). For an isolated square, for example, convexity and closure might be the cues that make us perceive the border as part of the square rather than the surrounding region (Fig. 1B). In the case of two overlapping figures, the occlusion features (T junctions) are generally thought to be important in establishing depth order between regions and thus border ownership (Fig. 1C). In random-dot stereograms (Julesz, 1960, 1971), the borders between regions of different disparities are always perceived as owned by the nearer surface.

We have seen that neural signals exist in the visual cortex that allow reliable discrimination of side of ownership, even without any averaging across cells (Fig. 18). The signals are fairly invariant against figure size (Fig. 19), which holds up even when large parts of the figures are cut off by the display margins, as in Figure 5. However, one could argue that the side preferences of single cells are just random variations of neural connectivity and have nothing

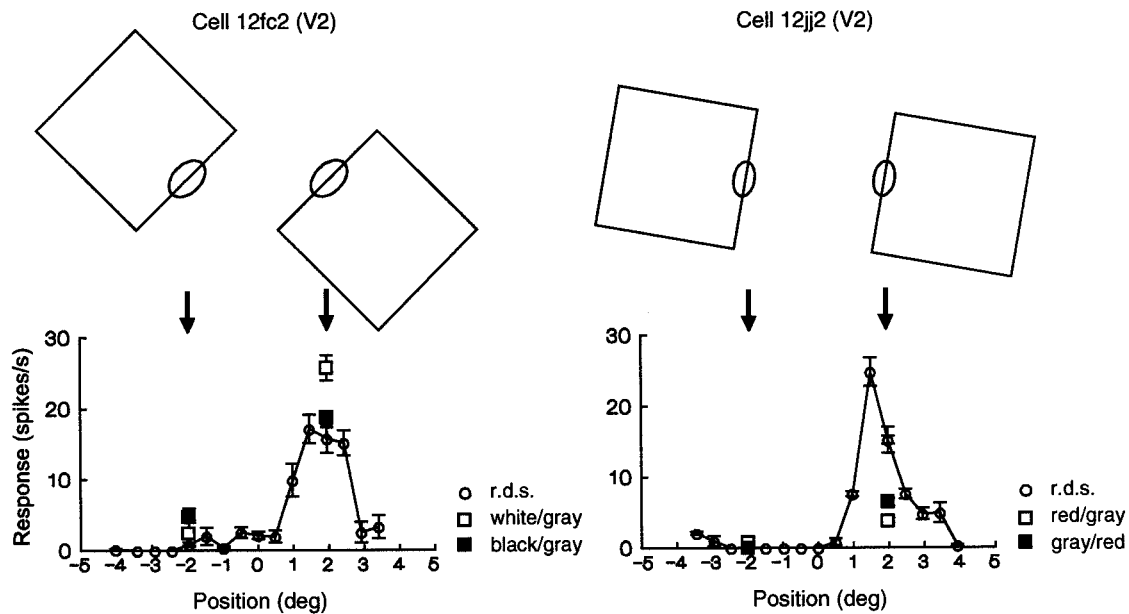


Figure 21. Border-ownership coding and stereoscopic edge selectivity. Responses of two cells of V2. *Circles connected by lines* show position-response curves obtained with random-dot stereograms (*r.d.s.*) portraying a 4° square. Disparity of square was set to optimum (7 arc min “near” for the left cell, 14 arc min “near” for right cell), and background disparity was zero. Position of square was varied orthogonally to preferred orientation. *Arrows* indicate positions at which edges were centered in the minimum response fields, as illustrated above. *Square symbols* represent responses to edges of uniform squares of two contrast polarities, as specified in legend. It can be seen that both cells responded selectively to one side of the square, and that the preferred sides were the same for contrast-defined and disparity-defined squares. Because random-dot stereograms define border ownership unequivocally, these results confirm the assumption that the side preferences for contrast-defined figures reflect border ownership coding. Location of receptive fields: left cell (2.9°, -1.6°), right cell (-3.2°, -3.1°).

to do with perception of border ownership. If the response modulations reflect mechanisms in object perception, as we hypothesize, they should hold up also when information about border ownership (depth order or figure-ground relationship) is provided by other cues. If the side preferences of single cells were just random variations, the preferred side of ownership obtained for isolated squares would often differ from the preferred side obtained with overlapping figures or edges in random-dot stereograms.

Another question that we seek to answer is that of the general type of mechanism that produces the side-dependent response modulation. Border-ownership discrimination requires considerable convergence of signals, and we have seen that some cells of V2 and V4 indeed integrate context as far as 10° from the receptive field center of the cell. The question of how the system deals with the problem of cue and context integration is generally discussed in terms of two classes of models. Both assume that basic mechanisms exist that evaluate the relevant cues such as contour shape, occlusion features, disparity, etc. In the first type of models, figure-ground segregation is a process of several stages with increasing degrees of complexity. The outputs of the basic mechanisms are successively combined, and image context is integrated stage by stage. This is the “bottom-up” model. The other type of models are feedback networks in which the ambiguous low-level information is integrated at a high level at which different mechanisms converge and image context is available, and this information is fed back to the lower stages. The network disambiguates information and resolves conflicts in an iterative process (“top-down” model). Bottom-up models predict that border-ownership signals show various forms of cue effectiveness and cue integration, depending on the mechanism to which the signal belongs and its level in the hierarchy. Signals at the low levels may also be contradictory if cues disagree. Except for time integration of afferent signals, no improvement of signals takes place, and no conflicts are resolved, over time. By contrast, top-down models predict that all signals are consistent once the network has settled into a stable state, even those at the low levels. (We consider here relatively simple displays in which figure-ground perception is generally clear and unambiguous.) At the beginning, signals might vary in the degree to which

they reflect the available cues, and possible inconsistencies might occur, but only during a brief period after stimulus onset. Thus, we have three alternative hypotheses with different predictions: (H0) Side selectivity occurs randomly. The side preferences of a cell for different types of displays will not be related to the perceived sides of ownership. (H1) Bottom-up model. Each cell may show side preference for some cues, but not others. If different cues are effective, the preferred sides will be consistent with perceived border ownership. (H2) Top-down model. The side preferences of each cell will be consistent with perceived border ownership across different displays (specifically those for which perception is unambiguous and clear).

Binocular disparity

Whereas two-dimensional displays are generally somewhat ambiguous (the white square in Fig. 1B can also be perceived as a window, the dark region in Fig. 1C can be perceived as a partly occluded object or as an overlain L-shaped object), stereoscopic displays can make perception unambiguous. As mentioned, edges in random-dot stereograms are always perceived as belonging to the nearer surface, just as edges in real space. In area V2, we have found cells that signaled edges in random-dot stereograms (von der Heydt et al., 2000). These cells were orientation-selective and usually also selective for the polarity of disparity edges. Although found only in a minority of cells of V2 (~1/6), stereoscopic edge selectivity is a conspicuous phenomenon that can be a tool in interpreting the function of neural signals. Figure 21 shows the results obtained from two stereoscopic edge cells that were tested for border ownership with contrast-defined, uniform squares (standard test). The circles connected by lines show the variation of responses when a stereoscopic square was presented at various positions relative to the response field of the cell (ellipse). The insets illustrate the two positions at which an edge was centered in the response field. Only one of these produced responses. Square symbols represent the responses obtained with uniform squares. It can be seen that the same side of figure was preferred for the uniform squares as for the random-dot stereograms. In the latter, the depth order of the edges is defined locally, by the disparities of

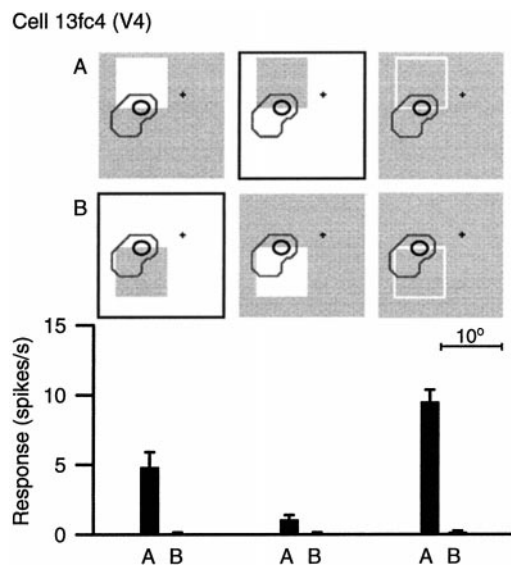


Figure 22. Consistent border-ownership coding for *solid* and *outlined* squares. Cell of V4. Conventions of receptive field map as in Figure 12. Colors were *gray* and *pink* (depicted as *white*). Line width, 0.2°.

the dots in the neighborhood, whereas the former provide no local clue as to which side is figure and which is ground. This result supports the assumption that the side selectivity revealed by the standard test is related to perception of border ownership. Since we changed to the large-field, direct-view system in later experiments we have not tested more border-ownership cells with stereoscopic displays.

Solid color versus outlined figures

Color certainly helps in defining objects; regions of the same color tend to be grouped in perception. In displays such as those of Figure 2, the system might define as an object a small area of color that is surrounded by a large area of different color. However, line drawings can also produce the perceptual differentiation of figure and ground. The contour of an outlined square is perceived as belonging to the enclosed region, even though this region has the same color as the surround. Figure 22 illustrates the responses of a border ownership-selective cell of V4 to solid and outlined squares. Figure 25 below shows another example of border-ownership selectivity with outlined figures. In total, we have tested 32 cells with outlined squares (most of which were selective for thin bars in the receptive field, but were unresponsive to contrast edges); 13 showed a significant effect of figure side (0 of 3 of V1; 2 of 11 of V2; 11 of 18 of V4). Thus, outlined figures were about as effective in producing border-ownership signals as solid color figures. Nine cells that were border ownership-selective in the standard test were also tested with outlined squares, and four of these showed significant consistent modulation, whereas five showed no significant modulation. This suggests that, in some cells, luminance and color differences (or the consistent border contrast at the contours) play a role in figure-ground differentiation. Remember that outlined figures were mostly tested in cells that responded much better to thin lines than edges. Further experiments are needed to clarify the role of surface color in border ownership coding.

Occlusion cues

Two overlapping figures were used to assess the effectiveness of occlusion cues (Fig. 3). Figure 23 shows an example of this test in columns 5 and 6 (the standard test is shown in columns 1 and 2). As the ellipses indicate, the contour between the two figures was centered on the receptive field. The arrangement of corners and T-junctions differed between *A* and *B*, whereas the overall distribution of colors was quite similar. The neuron whose responses are plotted at the bottom of Figure 23 was a border ownership-selective

cell recorded in area V2. The cell was also strongly color-selective, with a preference for violet (CSI = 0.79) and selective for the polarity of edge contrast, producing almost no response to the gray–violet edge (columns 2, 4, and 6). As can be seen, the responses to the violet–gray edge were stronger when the edge belonged to a violet square on the left than when it belonged to a gray square on the right, in the standard test (column 1), as well as in the two-figure test (column 5; $p < 0.001$). Figure 24 shows an example of a cell that signaled border ownership for both contrast polarities in the standard test (columns 1, 2) but failed to do so in the two-figure test (5, 6). This cell was recorded in V4 and was also color-selective (green; CSI = 0.73). Figure 25 shows an example of a line-selective cell of V4 that was tested with outlined squares and with overlapping outlined figures of various amounts of overlap. A response difference that was consistent with the single-figure result was obtained for the largest overlap (3.3°; $p < 0.001$; column 2), but smaller overlaps did not produce significant differences. Similar dependence on the size of overlap was also found in cell 13li1 of Figure 23 and in other cells. It might reflect the limits on resolution of detail in peripheral vision.

Of 64 cells tested, 24 were selective ($p < 0.01$) for side of border ownership in overlapping figures. Of the 42 cells that were border ownership-selective in the standard test, 20 showed consistent side preference for overlapping figures (2 of 8 of V1 and V1–V2 border, 8 of 16 of V2, 10 of 18 of V4), 21 cells showed no significant modulation, and one cell (V4) reversed the side preference relative to the standard test.

Conflicting cues

In search of simple mechanisms that might explain the side selectivity in the above tests, we wondered if cells might be sensitive to the corners at the ends of the edge that is placed in the receptive field; the orientation of the corners would indicate the side of the square. We have tested this hypothesis by presenting C-shaped figures, as shown in columns 3 and 4 of Figure 23. In these figures, the central edge and its terminating corners (the inner part of the C) were features that occurred in displays 1 and 2 as part of the squares. In *A3*, for example, they would indicate that a square is located on the bottom left side, whereas perceptually the edge belongs to the gray C figure on the opposite side. The bar graphs below, representing the responses of a cell of V2, show stronger responses to *B3* than to *A3* ($p < 0.001$), in accordance with perception, and contradicting the corner selectivity hypothesis. A similar result was obtained in the V4 cell of Figure 24, which was side-selective with either contrast polarity. More often we found cells that were border ownership-selective in the standard test, but not for the C-shaped figure. Figure 26 shows an example of such a cell recorded in V2. Whereas local contrast did not affect responses in the standard test (1, 2), the cell showed a clear preference for dark–light local contrast in the C-figure test. We do not have an explanation for this kind of behavior, which was also observed in other cells. In total, 14 cells with significant border-ownership effect in the standard test were studied with C-shaped figures. Four of them showed consistent border-ownership preference (0 of 1 of V1, 1 of 6 of V2, and 3 of 7 of V4), while 10 showed no significant border-ownership modulation with the C-figures. None reversed the side preference. Thus, the corners terminating the edge seemed to play a role, because 9 of 14 cells lost their border-ownership selectivity, but their influence was not strong enough to tip the balance in any of the 14 cells. This indicates that the neurons also use other cues that are common to squares and C-figures.

Figure 27 summarizes the results of the experiments with overlapping figures and C-shaped figures. Each dot represents a cell tested. The left and right columns represent nonsignificant and significant border-ownership effects, respectively, as determined with single squares. This is illustrated schematically at the top, where line thickness indicates response strength (left column, no difference, right column, preference for figure–right). The rows represent the results obtained with overlapping figures and

Cell 13li1 (V2)

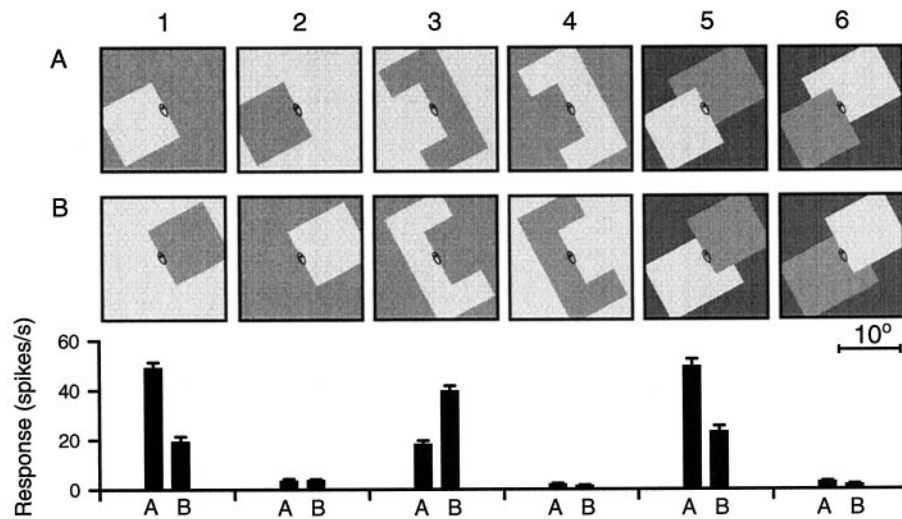


Figure 23. Example of a V2 cell tested with single squares, C-shaped figures, and overlapping figures. The cell was color-selective with a preference for violet. 1, 2, In the standard test the cell was found to be selective for border ownership and local contrast polarity, responding best to the edge of a violet square located on the bottom left-hand side of the receptive field (A1). 3, 4, With C-shaped displays, the cell responded better to B3, in which a violet C-shape was located on the bottom left, than to A3, in which the central portion was similar to A1. 5, 6, With overlapping figures, the cell responded to the violet-gray edge better when the violet figure appeared to be laid on top of a gray figure (A5) than when a gray figure appeared to be laid on top (B5). Location of receptive field (0.2° , -0.4°).

Cell 13md5 (V4)

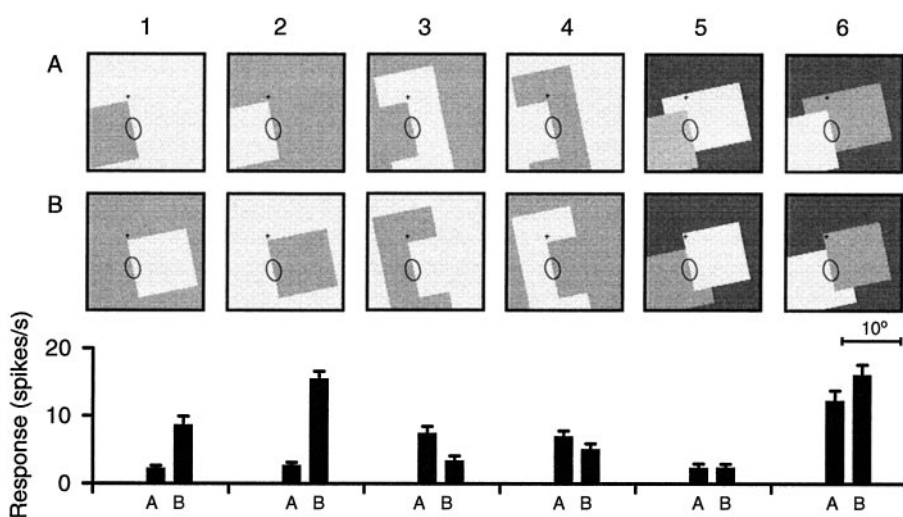


Figure 24. Example of a V4 cell tested with single squares, C-shaped figures, and overlapping figures. Ellipse indicates minimum response field. 1, 2, In the standard test, this cell showed a preference for figure location on the right (B) with both contrast polarities. The same side was preferred for C-figures (3A, 4A), but for the overlapping figures, response differences were not significant (5, 6). Colors: 1–4, green (depicted here as white) and light gray; 5, 6, green, light gray, and dark gray. Location of receptive field center (0.5° , -5.2°).

C-shaped figures, as illustrated on the left, based on a significance criterion of $p < 0.01$. It can be seen that cells that exhibited border-ownership selectivity with the single square (right column) nearly always preferred the same side of ownership with the other displays, if they showed any preference at all. Note also that 4 of 22 cells that did not differentiate sides with the single square did so with overlapping figures.

In summary, the test variations illustrated in Figures 21–26 demonstrate various degrees of cue invariance of neural border-ownership selectivity. We can exclude hypothesis H0, that side selectivity occurs randomly, because the side preference for different displays, if significant, varied in agreement with perception in almost every case. Hypothesis H2, postulating a cue-invariant representation of border ownership, can be ruled out because many cells were border ownership-selective in the standard test, but not for other displays, such as overlapping squares or C-shaped figures, in which border ownership is also clearly perceived. Our results are compatible with H1, which assumes that a variety of mechanisms for figure-ground discrimination exist in visual cortex, each of which uses a different computational strategy. If the outputs of these mechanisms are integrated within the same cortical areas, we expect to find signals with more or less complete cue invariance, depending on the level of integration, but no contradictory signals. That was what we found.

DISCUSSION

Cells of areas V1, V2, and V4 were studied by comparing the responses to contrast borders when these borders were displayed as part of different visual objects. This paradigm allowed us to measure the influence of the image context on local border responses and determine the range of image context that appears to be integrated in the neural responses. Slightly more than half of the edge-selective cells of V2 and V4 showed significant differences in firing rate depending on the side of the object to which the border belonged. A smaller fraction of cells of V1 also showed such an effect. In many cells the difference in firing rate was so strong that 1 sec of the activity of a pair of such cells would allow perfect discrimination of object side. The range of image context that was integrated in these neurons appeared to be larger than the 20° size of the displays that we used. The side-dependent differences emerged almost immediately after the onset of the responses and remained constant over several seconds during fixation. Experiments with displays in which perceptual border ownership was defined in different ways showed that shape cues, overlay cues, and binocular disparity are sometimes combined in single neurons to produce consistent border-ownership signals, but incomplete cue integration was the rule. We should point out that the various observations made in this paper are supported by unequal amounts of data. The significance of the effect of figure side in the standard

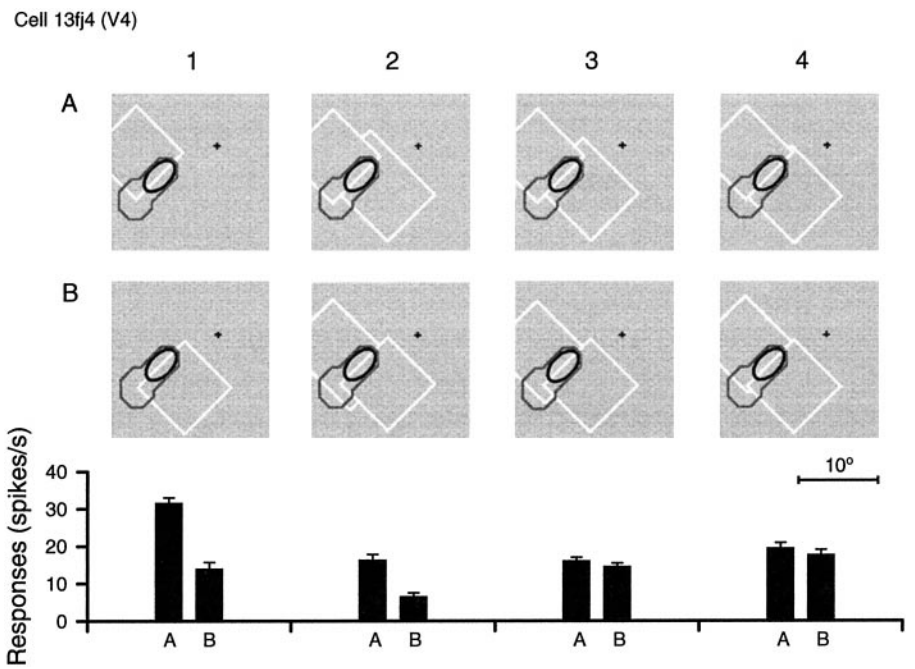


Figure 25. Responses of a cell of V4 to outlined squares and outlined overlapping figure displays. For the largest overlap ($1.7 \times 3.3^\circ$) the response difference was consistent with the side preference for single squares, but for smaller overlaps ($1.7 \times 1.7^\circ$ and $0.5 \times 0.5^\circ$) border ownership did not make a difference. White lines on gray, line width, 0.2° . Conventions of receptive field map as in Figure 12. Location of receptive field center ($-7.4^\circ, -3.9^\circ$).

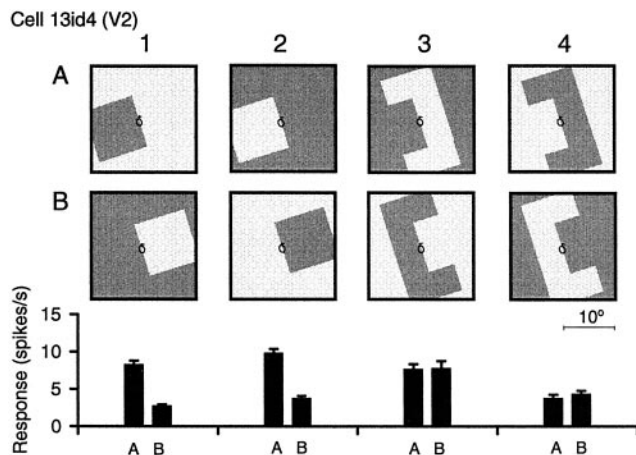


Figure 26. Responses of a cell of V2 (same cell as in Fig. 4). This cell showed contrast-independent side preference for edges of squares (1, 2). For C-figures, no side differentiation was found (3, 4). Instead, the responses now differed according to local contrast polarity. Colors gray and olive (depicted as dark gray) were used.

test and the reliability of the associated neural signals were determined in relatively large numbers of cells, but the various cues were studied only in smaller samples, and mostly in those cells that showed the border-ownership effect in the standard test. In particular, comparisons between solid and outlined figures and between stereoscopic and figural cues were made only in a few cases, and those examples of cue-invariant cells should be regarded merely as proofs of existence.

Before discussing the significance of the results in terms of border ownership coding we will consider several alternative explanations.

Average luminance and chromaticity

Our standard test consisted of a colored figure surrounded by a large region of different color (Fig. 2). Could the change in mean luminance or chromaticity between A and B (or C and D) account for the figure side-related response differences? The factorial analysis of the responses to the four displays of Figure 2, consisting of contrast-reversed pairs, separates out the effect of mean color (as discussed under Results, this effect appears as the interaction of

local contrast and figure side). Thus, the effect of figure side, as quantified in Figures 15, 16, and 18, is independent of the effect of mean luminance and chromaticity. Also, in 15% of the cells of V2 and V4 the responses showed a side-of-figure difference, but were

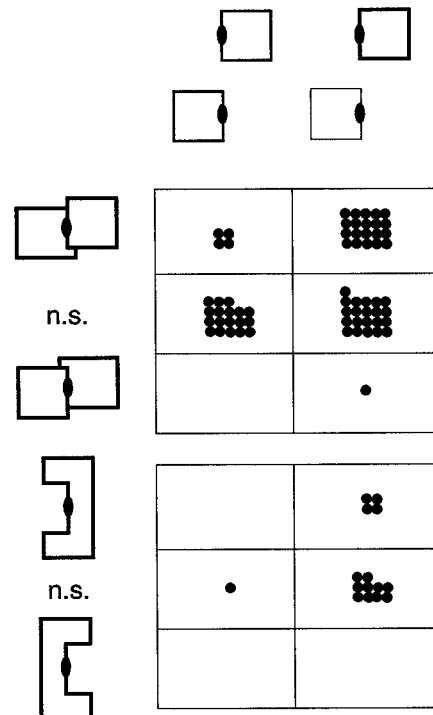


Figure 27. Summary of the results obtained with single squares, overlapping figures, and C-shaped figures. Each dot represents a cell tested. The right column represents cells with significant side preference for single squares, the left column represents cells with no significant preference. This is indicated schematically at the top, where line thickness represents response strength. The rows correspond to the side preferences in the other tests, as indicated on the left. If a significant ($p < 0.01$) side preference was found for overlapping figures, or C-figures, it was consistent with the single-figure result, except for one cell in which preference was reversed for overlapping figures. However, many cells with significant side preference for single squares failed to differentiate border ownership in the case of overlap or the concavity of the C, indicating incomplete use of the available cues.

not affected by the contrast polarity, indicating that the overall luminance–chromaticity had no influence. In other cells we found side-of-figure effects with outlined figure displays that are devoid of luminance and chromaticity differences. Furthermore, the size invariance (Fig. 19) shows that the figure side-related response differences did not depend on the presence of a small region of color surrounded by a large region of another color, because the figure could occupy half the display or more. Therefore, any simple mechanism based on the distribution of colors can be ruled out as an explanation.

Fixational eye movements

We have discussed this possibility already under Controls in the Results section. Eye movement recordings have shown that the differences in mean position of fixation between the presentations of the four different displays of the standard test were much smaller than the minimum response fields of most of the cells and did not depend systematically on the figure location (Fig. 10). Our analysis of the neural activity is based on responses sampled near the peaks of the position-response curves, where the response strength depends minimally on stimulus position (Figs. 11, 12), and any deviations of fixation could have produced only small random response differences. The adopted criterion of significance ($p < 0.01$) should have weeded out such differences. The reliability of the side-of-figure-related differences in repeated tests confirms this conclusion (Fig. 19). Thus, we are confident that even the side-of-figure effects found in some cells of V1 are real and not random fluctuations.

Location of the figure relative to fixation

We wondered if the preferred side of figure of the single neurons was related in any way to the location of its receptive field relative to the fovea. It is conceivable that responses were enhanced (or suppressed) for figures that were close to the fixation point, or overlapped it, compared to figures that were far from the fixation point. We have analyzed this relationship by dividing the directions around each receptive field center in four quadrants and counting the numbers of cells with the preferred figure location in the quadrant containing the fixation point, the quadrant opposite to the fixation point, and the other two quadrants. The percentages of cells in these three divisions were 23, 28, and 49%, respectively. Thus, the preferred figure side was not related to the location of the fovea relative to the receptive field. Although it is conceivable that the figure attracted attention, this cannot explain the side-related response differences, because the same figure location could be the preferred side for one cell and the nonpreferred side for another cell.

We interpret our findings as showing that, for each orientation and position in the visual field, the system has two sets of neurons to represent the two possible ways that a contrast border can be produced by an occluding contour in three-dimensional space. This coding scheme is illustrated in Figure 28. If the stimulus provides border-ownership cues, one member of each pair is activated more strongly than the other, as indicated by shading. Our results show that neural selectivity for border ownership is comparable to the known selectivity of cortical neurons for orientation, color, disparity, and direction of motion. Only a subset of cells in each of the areas that we have studied is selective for border ownership, and the degree of selectivity varies, just as for any of the other features. The extraction of any of these features requires certain kinds of image information whose availability varies from point to point and from image to image. What is different, of course, is that border ownership in general cannot be inferred from the local neighborhood of the border (the exception is the random-dot stereogram), whereas local cues usually work for the other features. To assign a contrast edge to one of the adjacent regions (the presumed foreground object) the system has to evaluate a global image context. Our finding that border ownership gradually emerges in the cortical hierarchy (Fig. 16) is consistent with this notion. Only a small proportion of cells was border ownership-selective in V1, whereas such selectivity was common in V2 and V4.

The influence of global figure-ground organization has recently

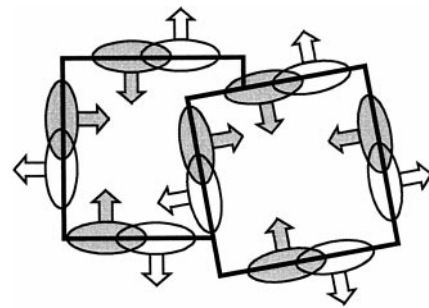


Figure 28. Schematic illustration of the cortical representation of contrast borders for a display of two overlapping squares. *Ellipses* represent location and orientation of receptive fields, *arrows* represent preferred side of ownership, and *shading* indicates activation of corresponding neuron. Each border is represented by two populations of orientation-selective cells whose relative activity codes the side of ownership. Thus, information about the location and orientation of borders, the color and luminance gradient at the border, and the side to which it belongs, is multiplexed in the responses of cortical neurons.

been demonstrated in responses of cells of V1 (Lamme, 1995; Zipser et al., 1996; Lee et al., 1998). Lamme (1995) used random textures of lines or dots in which figures were defined by a difference in orientation or motion of the texture elements. Responses were found to be enhanced when a figure was presented over the receptive field compared to a homogeneous texture. There are some differences in scope and design between those experiments and ours. Lamme (1995) and Zipser et al. (1996) used textured figures, and the figures were centered on the receptive field. Thus, an orientation-selective cell would respond according to the orientation of the texture, not the figure boundary. The interpretation was that, while the individual cells may signal qualities of local texture (or motion), the overall level of activity codes the figure-ground quality. The underlying coding hypothesis was that of an isomorphic representation of figure and ground regions. While this interpretation was convincing for textured figures, it left open the question of how figure-ground relations are coded for objects of uniform color, because it is known that most cortical cells respond only weakly, or not at all, to uniform stimuli (Hubel and Wiesel, 1968; see von der Heydt et al., 1996 for results in awake monkey). Lee et al. (1998) have also tested uniform figures with the same paradigm and found that, although the responses were weak, figure enhancement did occur. Nevertheless, the fact that the vast majority of cortical cells are orientation- and edge-selective shows that visual information is not represented in an isomorphic manner at this level. It seems implausible that figure-ground relations should be coded in the weak surface responses of neurons that respond vigorously to contrast borders. Our results show that figure-ground information is coded in the border responses. Thus, the same signals that encode form (by way of orientation selectivity) carry also border-ownership information.

Whereas we find border-ownership information only rarely represented in cells of V1, Lamme (1995) and Zipser et al. (1996) found figure enhancement in the great majority of cells of V1, and they emphasize the generality of this phenomenon. However, this point is not clear because their data included multiunit recordings. On the basis of multiunit activity, one cannot distinguish whether the response modulation is small but present in all cells, or large in some cells and absent in others. We found that a large proportion of cells in all the three cortical areas did not carry border-ownership information. In fact, the index of border-ownership discrimination reliability for V2 and V4 showed a tendency toward bimodal distribution (Fig. 18), suggesting a representation in which some cells carry this information, while others do not. Lee et al. (1998) have studied the responses to the borders of texture-defined stimuli, but do not comment on possible border ownership effects.

Size dependence

An important difference concerns the size dependence of the figure-ground influence. Zipser et al. (1996) found that the figure

enhancement dropped to zero for squares of 8–10° diameter. Our data of V2 and V4 indicate little diminution of border-ownership signals with square sizes up to 20° (Fig. 19). Also the few V1 cells tested showed almost no reduction up to 8–10°. This result agrees with human perception: informal observations with our displays indicated that perception of figure and ground was approximately equally strong for square sizes of 4 and 10°. The steep drop found by Zipser et al. (1996) disagrees with this observation. However, the size dependence of figure-ground perception might be different for texture-defined figures.

Time course

Another finding that differs between the studies is the time course of the figure-ground signals. Lamme (1995) found that, in the averaged signal, the figure enhancement appeared only after the initial peak of response, with a delay of 30–40 msec after response onset. In our data, the border-ownership difference appeared before the peak of the response and within 10–25 msec of the response onset (depending on the area). Our results impose stronger constraints on the possible mechanisms that might produce the figure-ground differences.

We should emphasize that our data on size dependence and time course are mainly from V2 and V4 (because border-ownership coding was rarely found in V1), whereas Lamme et al. recorded from V1; we have studied border responses, whereas Lamme et al. have studied surface responses. Thus, there is no contradiction between the results.

Using different configurations and modes of display (squares, C-shapes, overlapping rectangles, random-dot stereograms, and outlined figures) we found that many cells showed invariance to the cues that produce perception of border ownership. This indicates a remarkable convergence of different mechanisms. For example, the determination of depth order of edges in random-dot stereograms and the discrimination of figure side in displays of uniformly colored figures require different operations, and the utilization of overlay cues might require another completely different mechanism. Such invariance might indicate the presence of top-down signals, as envisioned in computational studies (Finkel and Edelman, 1989; Finkel and Sajda, 1992) (see also Lee et al., 1998). However, the frequency of neurons whose responses reflected only partial utilization of the available cues (Fig. 27) seems to contradict feedback models that aim to achieve a representation that integrates all these cues. It is true that we do not know the monkey's perception very well, and even human perception might be less well defined than we have assumed. However, what is significant in our results is that the combination of cues that are reflected in the signals differs between cells. For example, the responses of one cell, but not the other, might reflect overlay cues, whereas both signal border ownership for single-figure displays. Each cell has its answer to each of the displays presented, and this answer remains the same during a period of fixation, and over the 2 hr or so that we analyze its responses, but another cell has a somewhat different answer. However, perception of the displays does not change from one hour to the next, or from day to day. Thus, it is clear that neither V2 nor V4 provide the integrated representation of figure-ground information that corresponds to perception, which is what we would expect if the top-down model were correct. Our results suggest a mixture of parallel and hierarchical processes with predominantly bottom-up direction.

The currently prevailing view of the visual cortex as a hierarchy of stages of visual feature representation with gradually increasing size and complexity of receptive fields would attribute figure-ground segregation mostly to the higher levels, such as inferotemporal cortex, where cells have sufficiently large receptive fields. Finding global image context represented in cells of V2 (and even V1) is surprising. The degree of convergence of afferences up to this level (as evident from the sizes of the conventional receptive fields and their topographical scatter) is far too small to account for the extent of integration, and the intrinsic connections of V2 might be too slow for information to travel that far within the available

time. Using the formula for cortical magnification in V2 of Gattass et al. (1981), $M = 4 \text{ mm } E^{-1.1}$ (where E is retinal eccentricity in degrees) we calculate that, for the edge of a square of 4° size centered on a receptive field at 2° eccentricity, the cortical distance of the nearest corner of the square would be ≥ 2.5 mm, depending on the orientation of the square. Given a conduction velocity of horizontal fibers in monkey cortex of ~ 0.14 m/sec (Gonzalez-Burgos et al., 2000), the transmission would take at least 18 msec, plus the necessary synaptic delays. For larger figures the delay would be considerably longer. The border-ownership signals were not delayed by >25 msec, according to our measurements with figures of 4–6° size. Thus, intrinsic V2 mechanisms would be at their limits for these figures. We need to measure the delays for larger figures to decide whether such mechanisms could explain our results. Another possibility is that large-scale context integration occurs in V4, where some receptive fields are very large and retinotopy is less strict, and that backward projections from V4 contribute to border-ownership coding in V2. Back-projecting axons of V4 have terminal fields of 3–5 mm in V2 (Rockland et al., 1994). In terms of the recipient sites in V2 this spreading means additional convergence of topographic information. Thus, the V4–V2 projections might be well suited for context integration.

The importance of border ownership in perception has long been recognized (Rubin, 1921; Koffka, 1935; Nakayama and Shimojo, 1990; Driver and Baylis, 1996). By postulating one-sided assignment of borders to regions one can explain a wealth of observations on perception of figure and ground, depth order, illusory contours, brightness, color, and transparency (for review and model studies, see Grossberg and Mingolla, 1985; Finkel and Edelman, 1989; Nakayama et al., 1990; Nakayama and Shimojo, 1990, 1992; Finkel and Sajda, 1992; Anderson and Nakayama, 1994; Grossberg, 1997). Extrapolating our findings from monkey to human visual cortex, we propose that all these phenomena are based on the dimension of neural coding demonstrated in this paper. Our finding of joint coding of orientation, color, contrast polarity, and border ownership in cells of V2 and V4 fits well with those perceptual observations, but is at variance with the widely accepted view of separate feature-processing streams (Livingstone and Hubel, 1987; Zeki, 1991). Recent studies on cortical color and orientation selectivity have modified this view (Levitt et al., 1994; Leventhal et al., 1995; Gegenfurtner et al., 1996) (Zhou, Friedman, and von der Heydt, unpublished observations).

The discovery of neural border-ownership coding throws new light on the way visual information is represented in the cortex and could provide the clue to understanding visual feature binding and object representation. Our results add weight to the claim (Lamme, 1995) that global figure-ground processing occurs at low levels of the visual cortex. The recent demonstration that signals in visual cortex carry information about real object size (Dobbins et al., 1998) also indicates global processing and object-related coding. Together, these studies call for a reinterpretation of the function of the visual cortex.

REFERENCES

- Allman J, Miezin F, McGuinness E (1985) Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). *Perception* 14:105–126.
- Anderson BL, Nakayama K (1994) Toward a general theory of stereopsis: binocular matching, occluding contours, and fusion. *Psychol Rev* 101:414–445.
- Barlow HB, Blakemore C, Pettigrew JD (1967) The neural mechanism of binocular depth discrimination. *J Physiol (Lond)* 193:327–342.
- Baumann R, van der Zwan R, Peterhans E (1997) Figure-ground segregation at contours: a neural mechanism in the visual cortex of the alert monkey. *Eur J Neurosci* 9:1290–1303.
- Bullier J, Nowak LG (1995) Parallel versus serial processing: new vistas on the distributed organization of the visual system. *Curr Opin Neurobiol* 5:497–503.
- Dobbins AC, Fiser J, Allman JM (1998) Distance modulation of neural activity in the visual cortex. *Science* 281:552–555.
- Driver J, Baylis GC (1996) Edge-assignment and figure-ground segmentation in short-term visual matching. *Cognit Psychol* 31:248–306.
- Finkel LH, Edelman GM (1989) Integration of distributed cortical sys-

- tems by reentry: a computer simulation of interactive functionally segregated visual areas. *J Neurosci* 9:3188–3208.
- Finkel LH, Sajda P (1992) Object discrimination based on depth-from-occlusion. *Neural Comput* 4:901–921.
- Gattass R, Gross CG, Sandell JH (1981) Visual topography of V2 in the macaque. *J Comp Neurol* 201:519–539.
- Gegenfurtner KR, Kiper DC, Fenstemaker SB (1996) Processing of color, form, and motion in macaque area V2. *Vis Neurosci* 13:161–172.
- Gilbert CD, Wiesel TN (1990) The influence of contextual stimuli on the orientation selectivity of cells in primary visual cortex of the cat. *Vision Res* 30:1689–1701.
- Gonzalez-Burgos G, Barrionuevo G, Lewis DA (2000) Horizontal synaptic connections in monkey prefrontal cortex: an in vitro electrophysiological study. *Cereb Cortex* 10:82–92.
- Grossberg S (1997) Cortical dynamics of three-dimensional figure-ground perception of two-dimensional pictures. *Psychol Rev* 104:618–658.
- Grossberg S, Mingolla E (1985) Neural dynamics of form perception: boundary completion, illusory figures, and neon color spreading. *Psychol Rev* 92:173–211.
- Heitger F, von der Heydt R, Peterhans E, Rosenthaler L, Kübler O (1998) Simulation of neural contour mechanisms: representing anomalous contours. *Image Vision Comput* 16:409–423.
- Henry GH, Goodwin AW, Bishop PO (1978) Spatial summation of responses in receptive fields of single cells in cat striate cortex. *Exp Brain Res* 32:245–266.
- Hubel DH, Wiesel TN (1968) Receptive fields and functional architecture of monkey striate cortex. *J Physiol (Lond)* 195:215–243.
- Hubel DH, Wiesel TN (1977) Ferrier lecture: functional architecture of macaque monkey visual cortex. *Proc R Soc Lond B Biol Sci* 198:1–59.
- Julesz B (1960) Binocular depth perception of computer-generated patterns. *Bell Syst Tech J* 39:1125–1161.
- Julesz B (1971) Foundations of cyclopean perception. Chicago: University of Chicago.
- Knierim JJ, Van Essen DC (1992) Neuronal responses to static texture patterns in area-V1 of the alert macaque monkey. *J Neurophysiol* 67:961–980.
- Koffka K (1935) Principles of gestalt psychology. New York: Harcourt, Brace, and World.
- Lamme VAF (1995) The neurophysiology of figure-ground segregation in primary visual cortex. *J Neurosci* 15:1605–1615.
- Lee TS, Mumford D, Romero R, Lamme VAF (1998) The role of the primary visual cortex in higher level vision. *Vision Res* 38:2429–2454.
- Leventhal AG, Thompson KG, Liu D, Zhou Y, Ault SJ (1995) Concomitant sensitivity to orientation, direction, and color of cells in layers 2, 3, and 4 of monkey striate cortex. *J Neurosci* 15:1808–1818.
- Levitt JB, Kiper DC, Movshon JA (1994) Receptive fields and functional architecture of macaque V2. *J Neurophysiol* 71:2517–2542.
- Livingstone MS, Hubel DH (1987) Psychophysical evidence for separate channels for the perception of form, color, movement, and depth. *J Neurosci* 7:3416–3468.
- Nakayama K, Shimojo S (1990) da Vinci stereopsis—depth and subjective occluding contours from unpaired image points. *Vision Res* 30:1811–1825.
- Nakayama K, Shimojo S (1992) Experiencing and perceiving visual surfaces. *Science* 257:1357–1363.
- Nakayama K, Shimojo S, Ramachandran VS (1990) Transparency: relation to depth, subjective contours, luminance and neon color spreading. *Perception* 19:497–513.
- Nelson JJ, Frost BJ (1978) Orientation-selective inhibition from beyond the classic visual receptive field. *Brain Res* 139:359–365.
- Newsome WT, Britten KH, Movshon JA, Shadlen M (1989) Single neurons and the perception of visual motion. In: *Neural mechanisms of visual perception* (Lam DMK, Gilbert CD, eds), pp 171–198. The Woodlands, TX: Portfolio.
- Peterhans E, von der Heydt R (1989) Mechanisms of contour perception in monkey visual cortex. II. Contours bridging gaps. *J Neurosci* 9:1749–1763.
- Pettet MW, Gilbert CD (1992) Dynamic changes in receptive-field size in cat primary visual cortex. *Proc Natl Acad Sci USA* 89:8366–8370.
- Regan D, Beverley KI (1985) Postadaptation orientation discrimination. *J Opt Soc Am (A)* 2:147–55.
- Rockland KS, Saleem KS, Tanaka K (1994) Divergent feedback connections from areas V4 and TEO in the macaque. *Vis Neurosci* 11:579–600.
- Rubin E (1921) *Visuell wahrgenommene Figuren*. Copenhagen: Gyldendal.
- Schiller PH, Finlay BL, Volman SF (1976) Quantitative studies of single-cell properties in monkey striate cortex. I. Spatiotemporal organization of receptive fields. *J Neurophysiol* 39:1288–1319.
- Sillito AM, Jones HE (1996) Context-dependent interactions and visual processing in V1. *J Physiol (Paris)* 90:205–209.
- von der Heydt R, Peterhans E (1989) Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. *J Neurosci* 9:1731–1748.
- von der Heydt R, Peterhans E, Baumgartner G (1984) Illusory contours and cortical neuron responses. *Science* 224:1260–1262.
- von der Heydt R, Heitger F, Peterhans E (1993) Perception of occluding contours: Neural mechanisms and a computational model. *Biomed Res [Suppl]* 4:1–6.
- von der Heydt R, Zhou H, Friedman HS (1996) The coding of extended colored figures in monkey visual cortex. *Soc Neurosci Abstr* 22:951.
- von der Heydt R, Zhou H, Friedman HS (2000) Representation of stereoscopic edges in monkey visual cortex. *Vision Res* 40:1955–1967.
- Zeki S (1991) Functional specialization in the visual cortex: the generation of separate constructs and their multistage integration. In: *Signal and sense: local and global order in perceptual maps* (Edelman GM, Gall WE, Cowan WM, eds), pp 85–130. New York: Wiley.
- Zipser K, Lamme VAF, Schiller PH (1996) Contextual modulation in primary visual cortex. *J Neurosci* 16:7376–7389.