Spatial Distribution of Contextual Interactions in Primary Visual Cortex and in Visual Perception

MITESH K. KAPADIA,1,2 GERALD WESTHEIMER,1 AND CHARLES D. GILBERT1
1The Rockefeller University, New York, New York 10021; and 2Department of Neuroscience, University of Pennsylvania School of Medicine, Philadelphia, Pennsylvania 19104

Received 21 December 1999; accepted in final form 22 June 2000

Kapadia, Mitesh K., Gerald Westheimer, and Charles D. Gilbert. Spatial distribution of contextual interactions in primary visual cortex and in visual perception. J Neurophysiol 84: 2048–2062, 2000. To examine the role of primary visual cortex in visuospatial integration, we studied the spatial arrangement of contextual interactions in the response properties of neurons in primary visual cortex of alert monkeys and in human perception. We found a spatial segregation of opposing contextual interactions. At the level of cortical neurons, excitatory interactions were located along the ends of receptive fields, while inhibitory interactions were strongest along the orthogonal axis. Parallel psychophysical studies in human observers showed opposing contextual interactions surrounding a target line with a similar spatial distribution. The results suggest that V1 neurons can participate in multiple perceptual processes via spatially segregated and functionally distinct components of their receptive fields.

INTRODUCTION

An important task of the visual cortex is to integrate local information from different parts of a visual image into global percepts such as contours, surfaces, and three-dimensional shapes. Although the process of visuospatial integration has been traditionally ascribed to high-order cortical areas, there is a growing body of evidence that suggests that the primary visual cortex (V1) may play an important role. While the receptive fields (RFs) of V1 neurons, as measured by a simple stimulus, are quite small, stimuli outside of this region can have powerful modulatory influences when presented concurrently with stimuli inside the receptive field. The modulatory influences allow neurons to integrate information from large parts of the visual field and may allow neurons at this early stage of visual processing to participate in complex perceptual tasks such as contour integration and surface segmentation.

The existence of surround effects in V1 neurons, especially inhibitory effects, has been known for many years (Bishop et al. 1973; Gilbert 1977; Gulyas et al. 1987; Hubel and Wiesel 1965; Knierim and Van Essen 1992; Li and Li 1994; Maffei and Fiorentini 1976), but it is now clear that excitatory interactions may play an equally important role in neural responses. While excitatory contextual interactions have been postulated to play a role in contour integration and saliency (Field et al. 1993), inhibitory interactions are thought to be important in the segmentation of surfaces and textures (Knierim and Van Essen 1992). If opposing neural interactions are needed for different perceptual processes, the question arises as to how these processes can exist in V1 neurons without canceling each other out. One possibility is that excitatory and inhibitory interactions are present in the same neurons but are located at different parts of the RF. To test this hypothesis, we designed physiological and psychophysical experiments to map the RF surround.

METHODS

Physiology

Experiments were performed using two, alert macaque monkeys (Macaca mulatta). The monkeys were comfortably seated in a primate chair 1.5 m from a computer monitor with a resolution of 1,200 × 800 pixels refreshed at 60 Hz. Experiments were performed under photopic conditions with ambient light. All procedures complied with the National Institutes of Health Guide for the Care and Use of Laboratory Animals and were approved by the institutional animal review boards.

Experimental and surgical procedures were as previously described (Kapadia et al. 1995). Briefly, stimuli were presented in the near periphery while animals performed a foveal dimming task to help them maintain tight fixations. The animals received a juice reward if they held fixation within a 0.7–1.0° window and indicated the dimming by releasing a lever at the appropriate time. Neural recordings were obtained in 600-ms epochs. The stimulus was turned on 200 ms after the start of the recording and presented for 100 ms. Each fixation trial consisted of three to five recording epochs separated by at least 300 ms.

The time window used to calculate the evoked response from each recording site was adjusted according to that cell’s composite temporal response profile over the entire set of experiments performed on

Address for reprint requests: C. D. Gilbert, The Rockefeller University, 1230 York Ave., New York, NY 10021 (E-mail: gilbert@rockefeller.edu).

The costs of publication of this article were defrayed in part by the payment of page charges. The article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.
that cell. The number of spikes that occurred in this window was converted to spikes/second by dividing the response by the window length. Spontaneous activity for each experiment was defined as the neural activity in the 200 ms prior to stimulus onset, averaged across all stimuli within an experiment and converted to spikes/second. The displayed results represent the mean evoked response of the unit over the 8–15 trials of each stimulus after subtracting spontaneous activity. Error bars show one standard error of the mean above and below this value.

Neural activity was recorded from the operculum of striate cortex using glass-coated platinum-iridium electrodes. Neural signals were amplified, band-pass filtered between 300 and 3,000 Hz and fed through a time-amplitude window discriminator to isolate individual units and small clusters of two to three units. The results from the single-unit recordings were indistinguishable from the multi-unit recordings. All recording sites had complex cell-like properties and were from superficial cortical layers, the neurons of which form the main output to extrastriate visual areas.

At each recording site, we quickly estimated the basic properties of the unit under study such as its orientation tuning and RF center by displaying bars at different positions and orientations and assessing the neuron’s response on an audio monitor. The neuron’s optimal orientation was then determined quantitatively by recording neural responses to long oriented bars (measuring 120 × 3') spaced 10–20° apart over a full range of 180°. Recording sites that showed poor orientation tuning were omitted from the study. The RF center was determined by presenting small optimally oriented bars (measuring 15 × 3') at adjacent positions along the length and width axes of the RF. The center was defined as the location that produced the largest response. The remainder of the experiments, including the two-dimensional mapping studies, was done with bars measuring 30 × 3'. The length of the stimulus corresponds to the approximate size of RFs at this eccentricity. The stimuli were not scaled according to the dimensions of individual RFs. We use the word “target” to represent a stimulus placed at the center of a neuron’s RF and “flank” to represent peripheral stimuli.

The eccentricities of RF centers ranged from 2 to 7° from the fovea, averaging around 4°. The central target stimulus was always presented at the optimal orientation at the center of the units’ RFs. Since the experiments were performed over a narrow range of eccentricities, RF sizes tended to be similar to one another, and we chose not to scale the size of the stimuli between individual recording sites.

The stimulus set used to create the two-dimensional maps consisted of 82 unique stimuli interleaved in a random block format. Thus a minimum of 656 stimulus presentations (82 stimuli × 8 trials for each stimulus) were needed to create each set of maps under each set of contrast conditions. Since the subtraction techniques used to create the context and nonlinearity maps depended strongly on the response to the central target presented alone, we often interspersed an additional 8–32 presentations of this stimulus to minimize the uncertainty in its mean response.

Although the fixation window provided limits for continuing a trial, the animal maintained much more precise fixation than that allowed by the window. Eye positions were monitored and recorded at 100 Hz using a scleral search coil system (C-N-C Engineering). An analysis of eye position data for 33 experiments is shown in Fig. 1. The data were summarized by calculating the mean eye position during each 100-ms stimulus presentation for a given experiment and calculating the mean and standard deviation of the distribution. The variability in eye position between trials (measured as 1 SD of the mean) for each of the 33 experiments is shown in Fig. 1A, and the average of the 33 values in Fig. 1B. The standard deviation in eye position, averaged across all experiments, was 2.7 min of arc in the horizontal direction and 2.9 min in the vertical position. This is much smaller than the grain of the grids used in the two-dimensional receptive field maps (15–30').

FIG. 1. Eye position data. Animals maintained much more precise eye position during stimulus presentation than that allowed by the fixation window used as a criterion for accepting trials and providing rewards. A: standard deviation of eye position, along the horizontal and vertical axes, during stimulus presentation is shown for each of 33 experiments. B: average standard deviation in eye position for the 33 experiments shown in A. The mean variation in eye position was 2.7 min of arc in the horizontal direction and 2.9 min in the vertical direction.

Psychophysics

Psychophysical experiments were performed on human observers using a nulling technique to measure the perceived orientation of a vertical line when a pair of additional, tilted lines was presented simultaneously. In these experiments, we use the word “target” to refer to the stimulus that the subject is asked to make a judgement about and “flank” to refer to the additional, contextual lines presented in the surround. Stimuli were presented on a CRT monitor with a resolution of 1,024 × 768 pixels refreshed at a rate of 60 Hz. Stimulus presentation was controlled by a Matrox Millenium video card and a PC-compatible computer. No error feedback was provided. Observation was binocular with normal pupils and a free head. All human experiments were approved by the institutional committees on human experiments.

The foveal tilt illusion experiments were performed at an observation distance of 6 meters. The central target and flanks consisted of white vertical lines, 8 × 1' in size. The line length used in this and other psychophysical experiments were given values that approximate the size of V1 receptive fields. Unless otherwise noted, the Michelson contrast of the stimuli was 99%. All data points presented here are
based on at least 450 total trials for both flank orientations distributed over a minimum of 2 days.

Each trial consisted of a 2,000-ms cycle. The target and a pair of tilted flanks were presented for 100 ms and followed by a 1,900-ms interval during which the subject reported whether the target appeared tilted clockwise or counterclockwise by pressing the appropriate button on a computer mouse. The target line was accompanied by two flanks, which were symmetrically positioned...
with respect to the target. To concentrate on the orientation signal and to obviate any position clues, the stimulus was designed so that the ends of the flanks had the same relationship to the end of the target for all target and flank orientations. During each session, the location and orientation of the flanking lines remained constant but by changing flank orientations and positions from session to session a full range of values of these parameters could be explored.

During each presentation, the target was shown randomly at one of seven equally spaced orientations centered on vertical, and the observer reported, if necessary by guessing, whether it appeared tilted clockwise or not. A psychometric curve was fitted to the proportion of “yes” responses at the seven orientations by the method of probits (Finney 1952), where the mean value, i.e., the orientation at which “clockwise” and “counterclockwise” responses are equally probable, provides an estimate of the target orientation at which it has no apparent tilt (see Fig. 3). To eliminate possible biases, stimuli with clockwise and counterclockwise flanks were randomly interleaved in each experiment. The induced tilt, defined as half the difference between the means under conditions of clockwise and counterclockwise flanks, is a bias-free measure of the tilt illusion obtained by a nulling method based entirely on the observers’ “yes” or “no” responses to the question whether the target appeared to have a clockwise tilt in a minimum of 600 presentations with randomly distributed values of target line orientation and direction of flank tilt.

In general, there were no important qualitative differences between the fovea and the periphery in this class of psychophysical studies. To assure ourselves that the pattern of attractive and repulsive tilt illusions exists also in the periphery, one observer performed the experiments described above at an eccentricity of 4° (view distance: 1m, line lengths: 30’). Care was exercised to factor out asymmetries depending on retinal areas by accumulating data across eight meridia.

RESULTS

Two-dimensional contextual maps

PHYSIOLOGY. The physiological experiments were designed to study the spatial distribution of excitatory and inhibitory surround interactions around a neuron’s RF. After quantitatively mapping the center of the RF and finding the units’ optimal orientations, we presented combinations of stimuli inside and outside the RF. The total stimulus set was composed of 82 stimuli positioned in a 9 × 9 grid centered on the RF (Fig. 2A). The dimensions of this grid were fixed and were not scaled to the size of individual RFs. The stimulus set was designed to minimize the number of stimuli required to map the RF and its surround and therefore minimize the time required to maintain isolation of the recorded units. All stimuli were presented at the optimal orientation for each cell under study. The stimulus set consisted of conditions where either the central target was presented alone, the target was presented in conjunction with a pair of flanks located symmetrically with respect to the RF center, or the pair of flanks were presented in isolation.

The main findings of these experiments are depicted in Fig. 2B. The central target, presented in isolation, produced a response of 8.5 spikes/s. When collinear flanks were presented simultaneously, the response grew to more than three times the baseline level even though the flanks alone elicited little re-

FIG. 2. Construction of 2-dimensional physiological maps and main results. A: description of stimuli. In the physiological experiments we measured V1 neural responses to a stimulus set consisting of 82 unique stimuli that were combinations of an optimally oriented central target (blue box) and a pair of mirror-symmetric flanks that shared the same orientation as the center stimulus. The flanks were positioned at the center of each location within the 2-dimensional grid that is shown in the figure. Since a pair of flanks was used in each stimulus, the entire grid was covered by half of the shaded positions shown in the figure. The stimulus set consisted of 40 stimuli where the flanks were presented alone, 40 stimuli where the flanks were presented in conjunction with a center stimulus, 1 stimulus where the central line was presented in isolation, and another stimulus where 2 flanks were positioned side-by-side at the central position (used to construct RF map, C, top). The entire stimulus set was repeated between 8 and 16 times in a random block manner. The results are presented as the average neural response over multiple presentations. B: responses for different stimulus configurations. Left: stimuli; right: neural responses (in spikes/s). The center-alone stimulus produced a response of 8.5 ± 4.1 spikes/s, but in the presence of collinear flanks, the response jumped to 37.7 ± 4.7 spikes/s even though collinear flanks produced little response on their own (3.9 ± 2.6 spikes/s). Lateral flanks reduced the response to the center stimulus when presented simultaneously (4.7 ± 2.0 spikes/s) even though the lateral flanks produced a positive response when presented in isolation (11.6 ± 3.5 spikes/s). Stimulus contrast = 20%. Flank positions shown in the schematics correspond to the closest grid positions to the receptive field (RF) center for each configuration. C: 2-dimensional map construction. At the left of each map is a description of the stimuli used to determine the value for that map at each grid position. Note that the positions of the flanking stimuli in this schematic represent relative positions only. The actual calculations utilized flanking stimuli placed in the appropriate locations for each grid position. The flanking stimuli in the schematic are displayed in a different color, green, to stress this point and point out the difference with the stimuli shown in B, which represent fixed flank positions. At the right of each map is a color scale, where blue is used to represent excitatory responses and red to represent inhibitory responses. Each map is scaled to the maximum excitatory or inhibitory value for that map. The actual neural data were plotted at the center of each grid location and intermediate positions were given values by linear interpolation. There was no smoothing of the data. RF map: a map of the RF was constructed using the neural responses to the stimulus conditions in which the flanks were presented alone, without the central target. The map shows the positions within the 2-dimensional grid that could directly activate the neuron under study. The central position of the RF map is the response to a special stimulus where 2 flanks are positioned in a side-by-side arrangement (equivalent to a single flank with twice the width). This is done so that the stimuli in all positions contain the same total amount of light. Note that this method tends to overestimate the size of the RF compared with the more conventional minimum response field measurement since each stimulus consisted of a pair of lines instead of a single line. The response to the central target alone is shown as the “baseline” value at the top of the map. Three-bar map: this map is constructed from the stimulus conditions where the target and flanks are presented simultaneously, and shows that flanking stimuli far from the RF center can modulate the response. Context map: this map demonstrates the influence of flanking stimuli on the center response. The value of the map at each position is obtained by subtracting the response of the 3-bar conditions from the response to the center alone. For comparison purposes, the result is normalized by the center-alone response. The map shows that the closest flank positions along the collinear axis are strongly excitatory, while most other positions produce weak inhibition. Nonlinearity map: this map is obtained by subtracting the center-alone and the flanks-alone conditions from the response to 3-bar stimulus and is normalized by the center-alone response. Like the context map, it shows excitatory interactions for collinear stimuli, but there is an additional region of strong inhibition seen at the sides of the RF. This arises because both the center-alone and flanks-alone stimuli produce a response from the neuron, but when the stimuli are presented simultaneously, the response is much less than the sum of the individual components (see preceding text).
response. When flanks were presented side by side with the center stimulus, they inhibited the unit’s response below the baseline level even though when presented alone, they elicited a strong response. The influence of surround stimuli thus showed significant nonlinearities: the response to the three-bar stimulus could not be predicted from the sum of the responses to the component stimuli. The responses showed further that excitatory and inhibitory inputs from the RF surround were spatially segregated. Excitatory inputs were present along the neurons’ collinear axis, while inhibitory inputs were present at the sides of the RF.

We used the neural responses to the entire stimulus set to construct two-dimensional maps of the RF and its interaction with its surround (Fig. 2C). The stimulus conditions in which the flanks were presented in isolation were used to create the RF map. This technique tended to overestimate the size of the RF compared with more traditional measurements of RF size, such as the minimum response field (not shown) since each presentation consisted of a pair of lines instead of a single line. Nonetheless this technique produced a detailed map of the parts of the visual field from which action potentials could be directly elicited. Similar to traditional methods used to map the RF, there is a strong response at the RF center, which falls off rapidly in all directions.

The three-bar map shows the magnitude of the responses, plotted as a function of the flank position, when the central target and flanks were presented simultaneously. This map is shown here for completeness but is omitted for brevity in the rest of the paper. The strongest responses to the three-bar stimuli occurred when the flanking lines were presented along the collinear axis close to the RF center.

By subtracting the neural response in the three-bar condition at each grid position from the center-alone response, we constructed “context maps.” This map displays how the addition of flanking lines at each grid position influences the neuron’s response to the central target. To allow comparison of these maps across different cells, each map was normalized to the cells’ response to the central target presented in isolation. Facilitatory interactions were located along the collinear axis, while inhibitory interactions were strongest in discrete regions at the sides of the RF. The four lobes of facilitation and inhibition were superimposed on a larger, weak field of diffuse inhibition. The context map reflects the total contextual influence of the flanking stimuli without subdividing this influence into linear and nonlinear components.

We isolated the nonlinear part of the contextual interaction in “nonlinearity maps,” created by subtracting both the center-alone and flanks-alone responses from the three-bar response at each grid position. This map highlights the neural responses that could not be explained by simple addition of the responses of the unit to the stimulus components, flanks and target, acting alone. This map was also normalized to the center-alone response. Compared to the context map there was strong inhibition at the sides of the RF, while the collinear excitation appeared qualitatively similar. The difference between the context and nonlinearity maps at the sides of the RF arises because at this grid position, the flanks-alone condition elicited a positive response from the cell, but when the same flanks were presented in conjunction with the central target, their effect was inhibitory (see Fig. 2B).

![Figure 3](#) Spatial dependency of repulsive and attractive perceived tilt effects in human subjects. A nulling technique was used to measure the perceived orientation of a vertical line when a pair of tilted, flanking lines was presented simultaneously. The flanking lines caused the central target to appear tilted—the direction of the induced tilt depended on the position of the flanks. A and B: psychometric curves in 1 observer. During each stimulus presentation, the central target was presented at 1 of 7 orientations. The observer’s task was to report whether the target appeared tilted clockwise or counterclockwise relative to vertical. Data points represent the percentage of times the observer reported a clockwise tilt in the central target when it was at each of the 7 orientations. The perceived orientation of the target is defined as the orientation where the target appeared tilted in either direction an equal number of times (dotted line at 50% positive responses) and was estimated by fitting the data with a psychometric function. Each experiment involved using 2 psychometric curves derived from mirror-symmetric stimulus sets (see insets). The magnitude of the induced tilt is defined as half the difference in perceived orientation between the 2 stimulus conditions, i.e., half the distance between the psychometric curves along the dotted line. Target-to-flank separation = 8°. Orientation difference between target and flanks = 5°. A: lateral flanks. The target appeared tilted away from the orientation of the flanks. This is the direction of tilt in the classical tilt illusion. B: collinear flanks. The psychometric curve was shifted in the opposite direction, such that the target appeared tilted toward the flanks. We call this an “attractive” tilt illusion. C: average data from 5 subjects for both flank configurations. Attractive tilts are given positive values and repulsive tilts are given negative values.
PSYCHOPHYSICS. The spatial segregation of excitatory and inhibitory contextual interactions at the neural level led us to search for a similar dichotomy at the level of visual perception. One type of contextual interaction that has been studied extensively at the psychophysical level is the tilt illusion. In the classical description of this illusion, the presence of tilted, flanking lines causes a target to appear tilted in a direction opposite to the orientation of the flanking lines, which is a “repulsive” effect (Gibson and Radner 1937; Westheimer 1990). We suspected that if opposing contextual interactions were segregated in different positions around the receptive field, a perceptual effect observed when a context is placed along a target line’s axis might be reversed for contextual stimuli placed along the orthogonal axis. Furthermore plotting the direction and magnitude of induced tilts with the flanks in different spatial positions should produce similar two-dimensional maps of contextual interactions as those seen in the physiological experiments.

We studied the influence of contextual stimuli on the perception of an object’s orientation by measuring the perceived orientation of a vertical line (target) when a pair of tilted, flanking lines was presented simultaneously with the target. All stimuli were presented at high contrast. The flanking lines caused the target to appear tilted in a direction that depended on the spatial position of the flanking lines. When the flanking lines were positioned in a side-by-side arrangement with the central target, the target appeared tilted away from the orientation of the flanks (Fig. 3A). This is the conventional repulsive tilt illusion. However, when the flanking lines were positioned in a collinear or end-to-end arrangement, the target appeared tilted toward the flanks, which we call an attractive tilt illusion (Fig. 3B). The direction of induced tilts in the two conditions was consistent across the five subjects studied (Fig. 3C).

In the next experiments, we sought to characterize the dependency of induced tilts on the orientation of the flanks and their separation from the central target. Collinear flanks induced attractive tilts when the orientation difference between the target and flanks was 5–10° and repulsive tilts at larger orientation differences (Fig. 4A). The strongest attractive tilt occurred at an orientation difference of 5°. Lateral flanks produced repulsive tilts for most orientation differences, with the largest effect occurring at an orientation difference of 20° (Fig. 4B). A significant attractive tilt illusion was observed at an orientation difference of 75°. This “paradoxical” tilt illusion has been observed in previous studies (e.g., Westheimer 1990), but was only seen in the lateral case here. Results averaged across subjects are shown in Fig. 4C.

Because the main purpose of these experiments was to study the positional dependency of contextual interactions, we con-

![Fig. 4](image-url) Dependence of tilt magnitude on orientation difference and flank separation. A–C: orientation dependence. Induced tilts are plotted as a function of the orientation difference between target and flanks. Target-to-flank separation = 8°. A: collinear flanks. Attractive tilts were observed for angular differences of 10° or less, and repulsive tilts were observed for larger orientation differences. B: lateral flanks. Most flank orientations produced repulsive tilts with a maximum effect occurring with an orientation difference of 20°. A small attractive tilt was observed with an orientation difference of 75°. C: average across subjects. Data points from individual subjects in A and B were averaged at each stimulus condition. D–F: target-flank separation. Induced tilts are plotted as a function of the separation distance between target and flanks. D: collinear flanks. Attractive tilts are observed for separation distances of approximately 20° or less, with small, repulsive tilts observed as the separation distance is increased. The orientation difference between target and flanks was 5°. E: lateral flanks. Repulsive tilts are observed at most separation distances, with smaller tilts as the separation distance is increased. Two of 4 subjects showed attractive tilts with very small separation distances. The orientation difference between target and flanks was 5°. F: average across subjects.
centrated the remainder of the psychophysical experiments on exploring the spatial distribution of the opposing tilt effects with an orientation difference between target and surround of 5°. This corresponds to the smallest orientation difference at which we could observe consistent attractive and repulsive tilts, therefore using target and flanks with similar orientation and illustrating the shifting balance between the two directions of induced tilt. For collinear flanks, the strongest attractive tilts were observed when the target and flanks were closest together. As the separation distance was increased, the attractive tilt declined, crossed zero near 20° separation and became repulsive at larger separations (Fig. 4D).

For lateral flanks, induced tilts were predominantly repulsive and became smaller as target-flank separation was increased (Fig. 4E). Significant repulsive tilts were observed over distances at least as large as 30°. Two of the four subjects showed small attractive tilts when the target and flanks were close together. However, both of these subjects reported difficulty in dissociating the tilt of the three lines of the stimulus pattern as separate, a crowding effect, but this was never a problem in the collinear condition. Results averaged across subjects are shown in Fig. 4F.

To study in a systematic manner the distribution of flank positions that result in attractive and repulsive tilts, we measured induced tilts when the flanks were positioned at a number of locations along a two-dimensional grid (Fig. 5A) analogous to the physiological experiments. Attractive tilts were found in a limited area along the collinear axis (Fig. 5B). Repulsive tilts were strongest along an axis orthogonal to the target line orientation, and weaker effects were observed in a diffuse distribution surrounding the target line. Interestingly, there was a second region of repulsive tilts that were intermediate in strength that was offset and parallel to the collinear axis. Results from a second subject are shown in Fig. 5C. This subject showed a more diffuse pattern of repulsive tilts than the first subject but a similar pattern of attractive tilts.

**Fig. 5.** Two-dimensional maps of tilt effects. A: arrangement of stimuli. We studied the spatial structure of induced tilts by measuring the direction and magnitude of the tilt illusion as a function of flank position. Like the previous experiments, each stimulus involved a pair of tilted flanks, positioned in mirror symmetric locations relative to the vertical target (red box). The flanks were positioned at the centers of a 2-dimensional grid, where adjacent grid positions were separated by 4° along the x axis and 8° along the y axis. The flank positions where the tilt illusion was measured are shaded in gray, and the data were reflected to produce the full map. Since each experiment involved a pair of symmetric flanks, the actual experiments correspond to 29 unique flank positions or half of the shaded area. Thus the total information in each map is contained in 1 quadrant or one-fourth of the map, but reflects stimuli that encompass half the grid positions. The induced tilt at each position was calculated from either 450 (subject MK) or 600 (subject KP) psychophysical responses, corresponding to a total of 13,050 or 17,400 responses respectively for each map obtained. B: results from 2 subjects. The direction and magnitude of the induced tilt at each grid location is depicted through the color scale shown at the right. Attractive tilts are shown in blue and repulsive tilts in red. Tilt magnitude is depicted through the shade of these colors. Actual data are plotted at the center of each grid position with intermediate positions determined by linear interpolation. There was no smoothing of the data. Example stimuli along each axis are shown to the left and bottom of the map.
To compare the spatial scales of the psychophysical and neurophysiological results, we examined tilt effects in the near periphery in three subjects at an eccentricity of 4° (Fig. 6A), which is similar to the mean eccentricity of RFs in the physiological studies. The dependence of induced tilts on orientation and flank position was similar to the results in the fovea. We constructed a two-dimensional map in one subject similar to the maps in the fovea (Fig. 6C). While the basic pattern of attractive and repulsive tilts is the same as in the fovea, the effects are seen over much longer distances. For example, attractive tilts were observed over separation distances of 8–16° in the fovea but extended to separations >2° in the near periphery.

**Contrast dependency**

**PHYSIOLOGY.** In the next set of experiments, we examined the effects of changing stimulus contrast on the pattern of contextual interactions. Initially, the contrast of the target and flanks were changed simultaneously. Figure 7 shows the results of experiments performed at one recording site using stimuli at three different contrasts. The raw neural response to the five main stimulus configurations at each contrast is shown in Fig. 7A. At a contrast of 20%, the central target presented in isolation produced a small response from the cell; but this response was greatly enhanced by the addition of collinear flanks. When the central target was presented with side-by-side flanks, there was little change in the neuron’s firing rate. At a contrast of 50%, collinear flanks increased the target alone response by a much smaller amount, while side-by-side flanks were inhibitory. The experiments at 30% contrast show intermediate results. Both collinear facilitation and lateral inhibition were evident at this contrast.

The pattern of contextual interactions at different contrasts were examined further in Fig. 7B, which shows the results of the two-dimensional map analysis at all three contrasts. At the lowest contrast tested (20%), collinear excitatory interactions were quite strong. Inhibition was weak and seen only in the nonlinearity map. As the contrast was increased to 30%, excitatory interactions became weaker and inhibitory interactions became stronger. Both excitatory and inhibitory interactions were observed in the context map. Finally, as the contrast was increased to 50%, only very weak excitatory interactions were seen in the context map, and there was no excitation seen in the nonlinearity map. Consistent with previous results, neurons’ minimum response fields are larger at high contrasts than at low contrasts (Kapadia et al. 1999). Several trial experiments...
FIG. 7. Effect of stimulus contrast on contextual interactions at a single recording site. This figure shows contextual interactions at 3 different levels of stimulus contrast. In these experiments, the contrast of the target and flank stimuli were varied simultaneously. 

A: main stimulus configurations. The neural response at 1 recording site is shown for the 5 main stimulus configurations tested at 3 different contrasts. As contrast was increased, interactions tended to shift from excitatory to inhibitory. Flank positions shown in the schematics correspond to the closest grid positions to the RF center for each configuration.

B: 2-dimensional maps. At 20% contrast, strong excitatory interactions were evident along the collinear axis, while inhibitory interactions were relatively weak. At 30% contrast, the context map showed both collinear facilitation and lateral inhibition, although the facilitation was weaker than at 20% contrast. At 50% contrast, collinear facilitation was quite weak and evident only in the context map. There was diffuse inhibition all around the RF, but the inhibition was still strongest at the lateral position.
were done at higher stimulus contrasts, but there appeared to be little difference in the results as target contrast was increased above 50%. These results were typical of all the recording sites studied. The findings show that excitatory interactions predominate at low contrasts and inhibitory interactions predominate at high contrasts.

To examine the effects of stimulus contrast at the population level, we averaged across all cells studied at the same spatial scale to create composite maps that reflect the entire neural population (Fig. 8). The effects were remarkably constant from one recording site to the next, and the two-dimensional map averaged over the population closely matched the individual example shown in Fig. 7B. At low contrasts, there was strong excitation in the collinear position and weaker inhibition in the lateral position, while inhibitory interactions dominated at high contrasts.

In the next set of experiments, we studied the effects of stimulus contrast by varying the contrast of the target and flanks independently. The results from a single recording site are shown in Fig. 9A. When the target and flanks were either both at high contrast or both at low contrast, the results were similar to those in the previous figure. However, when the target was kept at a low contrast and the flanks were presented at high contrast, there was a shift in the position of maximal facilitation toward more distant parts of the RF surround. The position of optimal facilitation was shifted slightly from the collinear axis at this recording site (bright blue region in nonlinearity map). Note the larger scale at which these experiments were performed compared with the maps in the previous figures (see scale bar).

A similar trend was observed across the population of cells. Composite maps created by averaging the results of the six cells studied under these conditions are shown in Fig. 9B. Excitatory interactions were seen in a region located close to the collinear axis, but at a greater distance from the RF center and in a more diffuse pattern than seen in the other parts of this study. We tended to observe more variability between recording sites for this experiment with the peak facilitatory position often falling slightly off of the collinear axis.

**Psychophysics.** The separation distance between target and surround at which one could see the maximum attractive tilt effect could be changed by altering the relative contrast of the target and surround. When the contrast of the target was lowered to 15% and the surround maintained at high contrast, the optimum separation for obtaining the tilt effect was 20–24’, and the effect was maintained out to 32’ (Fig. 10A). This is quite different from the condition in which all three lines were at high contrast, where the maximal effect was obtained with an 8–10’ separation and disappeared beyond 16’ separation (Fig. 10B). In the same manner, for both the physiology and the psychophysics, manipulating contrast influenced the spatial extent over which contextual influences were exerted.

**Discussion**

The principal conclusion that comes from this work is that contextual influences, both at the physiological and psychophysical levels, are not uniform but rather are highly dependent on the spatial positioning of the surrounding stimuli relative to the receptive field or to the target. While there are somewhat contradictory studies on the sign (facilitatory vs. inhibitory) and character of contextual interactions, our findings show that the RF surround is composed of both excitatory and inhibitory regions and that individual neurons are capable of displaying both types of interactions. We show further that the balance...
between facilitation and inhibition changes in a stimulus-dependent fashion.

The two levels of analysis (physiology and psychophysics) show that interactions of opposing sign are found in lobes located along orthogonal axes. Excitatory neural interactions and attractive tilt illusions were observed along the orientation axis of line segments, while inhibitory interactions and repulsive tilt illusions were observed at the sides of lines. The physiological and psychophysical interactions show similar dependencies on contrast, such that as one changes the relative contrast of center and surround, the facilitatory surrounds and attractive tilt influences extend farther from the central line.

**FIG. 9.** Effect of a low contrast target with high contrast flanks. *A:* single recording site. This figure shows contextual interactions when the target and flanks were both at high or low contrasts, as in the previous 2 figures, and for a low contrast target with high contrast flanks. When the target and flanks both had a contrast of 15%, there was strong collinear facilitation close to the RF and weaker inhibition at the sides of the RF. When the contrast of the flanks was changed to 50%, while the target remains at low contrast, the position of maximal facilitation remains near the collinear axis but shifts to a further position from the RF center. When both the target and flanks are at high contrast, only inhibition was observed. *B:* population average. Interaction maps obtained under identical conditions were pooled across the 6 cells studied with a low contrast target and high contrast flanks. The regions of strongest facilitation are located near or on the collinear axis but at relatively long distances from the RF center.
One might go further to relate the psychophysical and neurophysiological parts of this study via a simple, population-coding model of orientation discrimination, which can explain how tilt illusions can arise from excitatory and inhibitory contextual interactions at the neural level (Fig. 11). The neural basis of the tilt illusion, as supported by population coding models, may be the pattern of contextual interactions shown here. Previous implementations of this model have suggested iso-orientation inhibition as a possible basis for repulsive interactions in the orientation domain (Gilbert and Wiesel 1990).

Physiological experiments support the feasibility of this class of models (Gilbert and Wiesel 1990; Knierim and Van Essen 1992; Li and Li 1994; Li et al. 2000; Nothdurft et al. 1999).

As indicated in the model, the orientation tuning curves of neurons are sufficiently broad that a single oriented line activates cells optimally tuned to a wide range of orientation preferences. In this model, each neuron acts as a labeled line, signaling the presence of a stimulus at its own preferred orientation. The perceived orientation of a stimulus is derived from the activity of all neurons that are activated by that line. In this model, at the neural level, attractive tilts arise from iso-orientation facilitation and repulsive tilts arise from iso-orientation inhibition (Fig. 11, B and C). The flanking lines facilitate or inhibit neighboring neurons with similar orientation preferences, which skew the population vector in manner that induces a tilt in one direction or another.

Whether the dominant influence of contextual interactions is a change in the height of orientation tuning curves, as used in this model, or also involves changes in the tuning curves’ peak orientation or width requires further study. Once a population coding model of orientation is accepted—and there is little choice in view of the broad orientation tuning of neurons and the high precision of orientation perception of contours—one can think of several other ways of implementing interactions to yield the tilt illusion, but to distinguish between them, requires detailed knowledge of the changes in parameters of orientation tuning curves. The model presented here lends plausibility to the idea that the opposing tilt effects can be directly mapped onto the excitatory and inhibitory contextual zones around the receptive field and provides a suggestion for why the patterns of interactions observed in both parts of the study were so similar.

In the physiological experiments, we found that the balance between excitation and inhibition in the RF was not fixed but was regulated in a dynamic, stimulus-dependent manner. Low contrast stimuli tended to invoke strong excitatory interactions and weak inhibition, while the opposite was true at high contrasts. Contrast did not seem to change the positioning of the excitatory and inhibitory subregions but instead seemed to alter their strength and dimensions. Excitatory contextual interactions were observed at both high and low contrasts as evident in the context maps in Fig. 7. At low contrasts, the excitatory drive came from outside the neuron’s receptive field; the neuron’s response to the target and flanks presented simultaneously was more than the sum of the responses to the individual stimuli. At high contrasts, the excitation came from within the neuron’s minimum response field; the neuron’s response to the target and flanks presented simultaneously was more than the response to the target alone but less than the sum of the responses to the individual stimuli. Thus high and low contrast stimuli lead to a similar pattern of excitatory interactions in the context maps, but the nonlinearity maps look quite different under the two stimulus conditions. Also, while strong excitatory interactions were only observed close to the RF when the contrast of the flanks was low, excitation could be seen over much longer distances for higher contrast flanks.

Inhibitory contextual interactions were strongest at high contrasts. Under these conditions, there were lobes of strong inhibition along the orientation and orthogonal axes of the RF, which are consistent with the well known properties of end-inhibition (Hubel and Wiesel 1965) and side-band inhibition (Bishop et al. 1973), respectively. The current results lend further credence to the idea that end-inhibition is contrast dependent. High contrast stimuli invoke strong end-inhibition, while at low contrasts, the same regions at the ends of RFs are excitatory and little or no end-inhibition is observed (Kapadia et al. 1999; Sceniak et al. 1999).

While stimulus contrast is one factor that can influence the strength and sign of contextual interactions, there are likely to be many others. One factor that has been implicated in other studies is the complexity of the environment in which a stimulus is presented (Kapadia et al. 1995, 1999). When a high contrast stimulus is embedded in a complex surround, the neural response to that stimulus is often much less than the response to the same stimulus presented in isolation. The suppression induced by the complex surround also alters the...
contextual interactions are likely to be prominent at all levels of stimulus contrast not just at the relatively low contrasts found using the simple stimuli used here.

A strong candidate for the anatomical substrate of the contextual interactions observed here is likely to be the intrinsic, long-range horizontal connections formed by pyramidal cells in V1 (Gilbert and Wiesel 1979, 1983; Martin and Whitteridge 1984; Rockland and Lund 1982), although feedback from extrastriate areas may also play a role. Long-range horizontal connections can extend over distances as large as 6–8 mm and tend to connect cells with similar orientation preferences (Bosking et al. 1997; Gilbert and Wiesel 1989; Kisvárday et al. 1997; Ts’o et al. 1986). The extent of long-range connections at the anatomical level correlates well with the extent of contextual interactions observed in the psychophysical and physiological experiments. At an eccentricity of 4°, the average eccentricity of the RFs in this study, the cortical magnification factor is 2.5 mm/° (Dow et al. 1981). This means that horizontal connections can integrate information over 2–3° of visual space, which is similar to the extent of interactions we observed here.

Another important aspect of the long-range horizontal connections is their ability to provide both excitatory and inhibitory inputs to their postsynaptic neurons. Since the long-range connections arise from glutamatergic pyramidal neurons, they provide excitatory input directly but can also exert inhibitory effects through a disynaptic circuit involving inhibitory interneurons (McGuire et al. 1991). The dynamic change in long-range inputs from excitatory to inhibitory as a function of stimulus contrast is reminiscent of similar effects observed in intracellular recordings in cortical slices, where synaptic potentials evoked by a stimulating electrode change from excitatory to inhibitory depending on the intensity of electrical stimulation (Hirsch and Gilbert 1991; Weliky et al. 1995). In these experiments, low intensity stimulation evoked excitatory synaptic potentials, while higher intensity stimulation also activated inhibitory interneurons, resulting in strong, negative synaptic potentials that overwhelmed the smaller, excitatory potentials. Changing the contrast of a visual stimulus is likely to invoke similar cortical mechanisms as changing the intensity of an electrical stimulus.

The experiments in this study help delineate the basic structure of contextual interactions but likely reveal only a component of the full complexity of the interactions. The stimulus set used here was chosen to minimize the number of stimuli because of the difficulty in maintaining stable neural recordings in alert animals for long periods of time. For example, the stimulus imposes a mirror symmetry on the data since each experiment uses a pair of flanks instead of a single flank. Experiments with single flanks would be able to reveal asymmetries in the distribution of contextual interactions in individual cells. Also, the influence of a surround stimulus on a central target is dependent on the presence of additional contextual elements (Kapadia et al. 1995, 1999), and these cascading nonlinearities may explain higher-order aspects of visual processing. The experiments in this study were also limited to iso-orientation interactions; to characterize fully the role of V1 cells in analyzing complex visual scenes, one must study interactions between stimuli that differ in orientation.

A simple, conceptual model of how the observed pattern of contextual interactions for the stimulus. In effect, the suppression makes the neuron behave as if the central stimulus is at a lower contrast and excitatory contextual interactions become prominent at even the highest contrasts tested. This finding suggests that under real-world conditions, where objects are likely to be embedded in complex scenes, excitatory
contextual interactions might form the neural basis contour integration and surface segmentation is shown in Fig. 12. In this model, the saliency of a stimulus element results from the neural response of that feature relative to other elements in the surround. Collinear, excitatory interactions enhance the neural activity of stimulus elements that form smooth contours, resulting in enhanced saliency of these elements relative to other features in their surrounds. Lateral inhibitory interactions suppress the neural response of stimulus elements whose neighbors have the same orientation. The loss of this suppression in areas where there is a change in orientation results in enhanced saliency of the texture boundary.

The results of the current study suggest that the contextual influences may allow cells at early stages in cortical visual processing to mediate complex processes in intermediate level vision. The spatial segregation of excitatory and inhibitory inputs may allow individual V1 neurons to participate in multiple perceptual tasks that require opposing neural interactions. Our results suggest further that there is no clear separation between the stages of visual processing that serve for the analysis of simple

FIG. 12. Role of contextual interactions in contour integration and surface segmentation. A: comparison of psychophysical and physiological results. Two-dimensional maps obtained from psychophysical studies and from neural recordings are shown side-by-side for comparison. Both maps were obtained with high contrast stimuli. Psychophysical map: 5° flanks, fovea, subject MK. Physiological map: context map, average of 12 recording sites, 50% contrast. The basic features of the maps are quite similar. B and C: theoretical effects of contextual interactions on complex visual stimuli. Left: the input to the cortex; right: the output to the cortex. The darkness of each line is proportional to the neural response to that stimulus element. The model assumes that both the excitatory and inhibitory interactions are orientation tuned. B: iso-orientation facilitation and contour integration. In this type of contextual interaction, stimuli with similar orientations facilitate each other’s neural responses. This leads to an enhanced neural representation of the contour relative to its surroundings, enabling the contour to stand out from its environment. C: iso-orientation inhibition and surface segmentation. In this type of contextual interaction, line elements with similar orientations inhibit each other. This causes a suppression of neural responses in areas of uniform texture; but this suppression is lost at texture boundaries where there is a change in orientation. This pattern of interactions leads to enhanced neural activity at the texture borders relative to the surrounding areas and enables the segmentation of textures and surfaces defined by differences in orientation. The segregation of excitatory and inhibitory inputs allows the opposing interactions to exist simultaneously at the cellular level.
stimulus attributes, such as orientation, and those involved in higher order mechanisms of visuospatial integration.

We thank S. Kane for help with eye coil implantation and A. Glatz and J. Lopez for expert technical assistance.

This work was supported by National Institutes of Health Grants EY-07968 to C. D. Gilbert and MH-11394 to M. K. Kapadia.

REFERENCES


