The Role of the Primary Visual Cortex in Higher Level Vision

PERGAMON

Vision Research 38 (1998) 2429–2454

Vision Research

The role of the primary visual cortex in higher level vision

Tai Sing Lee [a,*], David Mumford [b], Richard Romero [a], Victor A.F. Lamme [c]

[a] *Center for the Neural Basis of Cognition and Department of Computer Science, Carnegie Mellon University, Pittsburgh, PA 15213, USA*
[b] *Department of Applied Mathematics, Brown University, Providence, RI 02912, USA*
[c] *Department of Medical Physics, University of Amsterdam and The Netherlands Ophthalmic Research Institute, 1100 AC Amsterdam, The Netherlands*

Received 14 January 1997; received in revised form 12 October 1997

Abstract

In the classical feed-forward, modular view of visual processing, the primary visual cortex (area V1) is a module that serves to extract local features such as edges and bars. Representation and recognition of objects are thought to be functions of higher extrastriate cortical areas. This paper presents neurophysiological data that show the later part of V1 neurons' responses reflecting higher order perceptual computations related to Ullman's (Cognition 1984;18:97–159) visual routines and Marr's (Vision NJ: Freeman 1982) full primal sketch, $2\frac{1}{2}$D sketch and 3D model. Based on theoretical reasoning and the experimental evidence, we propose a possible reinterpretation of the functional role of V1. In this framework, because of V1 neurons' precise encoding of orientation and spatial information, higher level perceptual computations and representations that involve high resolution details, fine geometry and spatial precision would necessarily involve V1 and be reflected in the later part of its neurons' activities.

Keywords: Figure-ground segregation; Medial axis transform; Primary visual cortex; Awake monkey electrophysiology; Non-classical receptive field

1. Introduction

David Marr's model for the computation of the meaning of images has dominated theory and experimentation for the last 20 years. In his influential book *Vision*, he proposed a series of computational modules representing steps in the analysis of an image and a rough correspondence between these modules and areas in cortex. Subsequent theoretical and experimental work refined his analysis, sometimes modifying it, sometimes making it more precise, but still following the basic ideas. For instance, middle temporal area (area MT) is considered to be the place where the aperture problem is solved, V2 the place where many gestalt grouping operations are performed. A central tenet of this model, however, is the decomposition of visual processing into successive feed-forward steps, into low, intermediate and high level stages, and the belief that visual cortex could likewise be divided into areas occupied with low, intermediate and high level operations. A key example was his strong assertion that the stereo correspondence problem could be solved before object recognition took place based on the psychophysical demonstration of human ability to see 3D structure in random dot stereo-grams. In Marr's framework, the primary visual cortex (area V1) is the site of the primal sketch, where local features are detected and grouped together into symbolic tokens and contours. He proposed a 2.5D sketch for the representation of surfaces and depth, and a 3D model a hierarchical modular representation of objects based on principle axes, as the bases of object recognition. These representations were thought to be computed and represented in the extrastriate cortices such as V4, MT and IT. Marr recognized that not all computations were exclusively feed-forward, although he seemed to believe that low-level vision can be done independently of later stages.

The purpose of this paper is to argue first on theoretical grounds that the low level visual computation cannot be completed before high level computations are begun; second, to present and examine neurophysiolog-

* Corresponding author. Tel.: +1 412 2681060; fax: +1 412 2685060; e-mail: tai@cnbc.cmu.edu.

0042-6989/98/$19.00
PII: S0042-6989(97)00464-1

ical evidence that V1 is computing different types of information during the 40–350 ms post-stimulus time period; and thirdly, to interpret these findings as a part of an extrastriate/striate feedback loop in which V1 plays a highly specific role much richer than simply carrying out the earliest stages of visual processing.

2. Conjectures on the role of V1 in visual processing

Object recognition in complex real world environments under multiple occlusions, perspective and lighting conditions is a very difficult problem. Before recognizing the object, it is often hard to segregate it from the background because on the one hand, its true contours are confused with local contrast edges caused by shadows, specularities, and surface discontinuities, and, on the other hand, the true object edges can be irregular, faint and partially occluded. To find the boundary of an object, the first set of contrast edges must be discounted and the second set must be enhanced or interpolated. But an object must be segregated from the background and its boundaries defined before one can compute its shape properties. These shape properties will have to be modified if the object is partly occluded or in shadow. An example is shown in Fig. 1: although the figure of the old man is extremely obvious to human perception, application of the popular Canny edge detector makes mistakes in all of the above. We believe that this figure cannot be separated from the background without substantial reconstruction of the 3D structure and the illumination of the scene. Curiously, the most recognizable object in the scene is the man's ear, which might, for example, entrain the process of matching next the face, and finally the body. In other words, figure-ground segregation and object recognition are intertwined: they cannot progress in a simple bottom-up serial fashion, but have to happen concurrently and interactively in constant feed-forward and feedback loops that involve the entire hierarchical circuit in the visual system. The idea that various levels in cognitive and sensory systems have to work together interactively and concurrently had been proposed in more general computational terms, particular by McClelland and Rumelhart [2] in terms of interactive activation neural networks, by Grossberg [3] in terms of adaptive resonance theory, by Mumford [4,5] in terms of pattern theory, by Ullman [6] in terms of counter-streams model, and Dayan et al [7] in terms of the Helmholtz machine.

If this hypothesis is true, one would expect to find that cells in V1 respond in very different ways in the initial phase of visual processing, e.g. 40–60 ms post-saccade or post-stimulus in the experimental situation, and in later stages, e.g. 60–200 ms post-saccade or post-stimulus. What effects would the computational theory lead one to expect? The analysis of V1 responses to basic elementary stimuli, such as bars and gratings, have suggested that early responses can be modeled as linear and nonlinear local filter responses, such as Gabor filters and the sum of squares of matched even and odd Gabor filters [8,9]. These, of course, detect all contrast edges and strips and respond well to many types of texture. Some of these will indicate object boundaries; others result from a multitude of illumination effects and surface properties of objects.

We should anticipate, therefore, that in some way the local edge contrast response will evolve, some increasing, some decreasing. Perhaps second order texture edges will be detected, illumination edges discounted, and faint edges that are highly significant for purposes of recognition will get enhanced. One would also expect other important figure/ground clues, such as T-junctions and illumination clues to modulate V1 responses. We can look for effects indicating that regions are being labelled, 'colored' in Ullman's terminology , which is a prerequisite to compute the shape of a region. Such responses might take the form of modula-

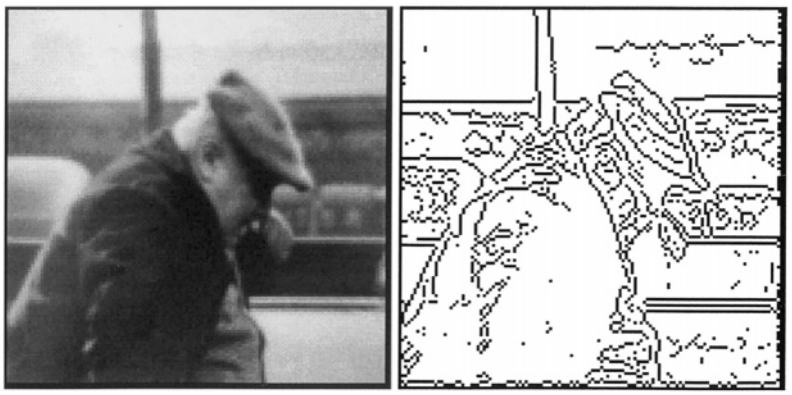

Fig. 1. An image of an old man and the edge signals produced by applying the popular Canny edge detector to the image. It illustrates that bottom-up edge signals are inherently difficult to interpret because of the ambiguities in local contrast edge information. Segmentation and recognition are necessarily intertwined, involving the entire hierarchical circuit of the visual system at the same time.

tions of texture responses which are describing surface properties.

A more radical theory proposes that tracing curves in images has a second, quite distinct use in visual processing. One of the main approaches to object recognition is the grammatical or structural approach, which is based on decomposing an object into primitive parts [10–14] and linking them together using a hierarchical framework like the parse tree of a sentence. This approach was, in fact, favored by Marr in his 3D model. It is particularly useful in encoding the great variety of complex biological forms. Biological bodies with flexible joints can change drastically with view point changes and motion. Blum [10] observed that under such changes, a region-based description based on the skeletons of the objects is much more stable than a boundary based description. He proposed that complex biological forms could be described efficiently using the skeletons and a small finite set of shape primitives. His skeleton, called the medial axis transform, is formally defined as the locus of centers of the largest circles inside the region, hence touching its boundary in two distinct points. Given that the medial axis, like boundary, involves curve tracing, which requires precise spatial precision and orientation resolution in a 2D topological map provided only by V1, one would expect that V1 neurons should be involved in the computation of these region and shape descriptors, and that the signals should be reflected in the later part of their responses.

If so much is being calculated in V1, is there any model to suggest which computations involve V1 and which do not? For instance, V1 neurons have not been found to solve the aperture problem or to respond to illusory contours. We believe that the results described here are consistent with the following revised model for the role of V1: that V1 is a unique high resolution buffer available to cortex for calculations, and will be used by any computation, high or low level, which requires high resolution image details and spatial precision. For example, why does the medial axis need high resolution? One reason is that we are extremely sensitive to symmetry and aspect ratio (for instance, it is heavily used to distinguish faces). A 10% change in aspect ratio makes a shape look very different. Only in a cortical area, where neurons are sensitive to disks of different diameters, can one compute the medial axis and the aspect ratio. Another much simpler reason for going back to the high resolution version of the stimulus is simply that some details that are overlooked as noise in the first pass often turn out to be crucial in confirming the identity of an object. In contrast, what V1 doesn't have are the collaterals that span the full width of the image. In computer vision, the image pyramid is computed precisely to make it possible to rapidly integrate information over the whole image. Such an image pyramid might be constructed in V1, V2, and V4 as Olshausen et al [15] pointed out in their proposal for how a window of attention at four different scales might be expected for computation in area IT.

In this paper, we report evidence from single unit recording in V1 of awake behaving macaque monkeys that lends support to the above conjecture. The evidence suggests that V1 seems to be involved in various higher order perceptual computations including the computation of cue-invariant or 'symbolic' surface boundaries, figure-ground (or inside-outside) distinction and the medial axis of shape. These computations have been described conceptually by Ullman [1] as visual routines, the processing of visual information beyond the creation of the early representations. These routines suppose to establish abstract shape properties and spatial relations that are vital to object recognition but are not represented explicitly in the initial representations of the visible environment.

3. Neurophysiological experiments

3.1. Background and motivation

It has been known for 20 years that neurons in area V1 are sensitive not just to the local features within their receptive fields, but are strongly influenced by the context of the surround stimuli. Many investigators have studied these surround or contextual modulations [16–28]. These contextual interactions have been shown to exert both facilitatory and inhibitory effects from outside the classical receptive fields. Both types of interactions can affect the same unit, depending on various stimulus parameters. Recent cortical models by Stemmler et al. [29] and Somers et al. [30] described the action of the surround as a function of the relative contrast between the center stimulus and the surround stimulus. These mechanisms are thought to mediate such psychological effects as filling-in [24] and pop-out [23].

It is in this context that we find Lamme's [25] findings and their interpretation to be particularly provocative and interesting. Lamme compared the responses of V1 neurons when their receptive fields were placed inside a texture figure, and when they were placed outside a figure in the texture background. He found that V1 neurons not only responded better inside the figure, but the enhancement was spatially uniform within the figure. Moreover, this enhancement was observed not only for texture cues, but also for motion cues. Based on these observations, Lamme [25] suggested that this interior enhancement effect might be related to a more abstract perceptual construct called figure-ground, a signal that conveys whether the neu-

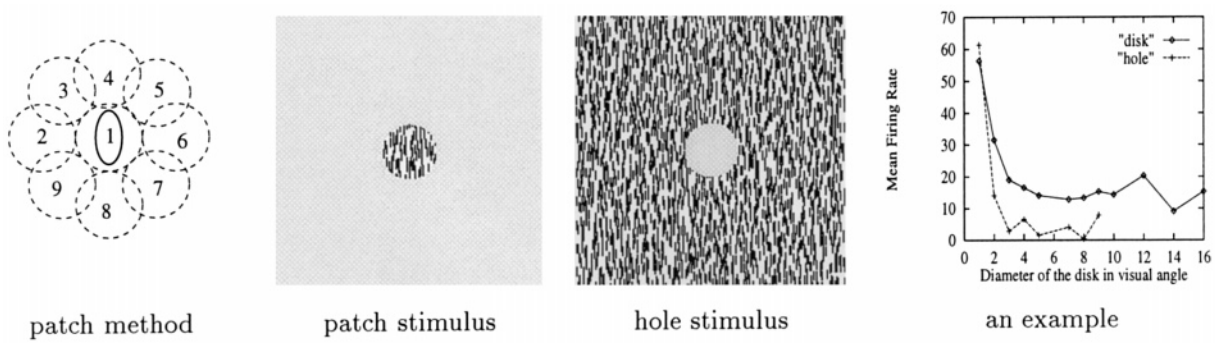

Fig. 2. The classical receptive field of each cell was first localized and its spatial extent ascertained by moving a black bar in different directions over the receptive field. Its maximum spatial extent was further assessed by successively flashing circular texture patches at and around the localized receptive fields (patch mapping method). The diameter of the smallest disk that elicited response at center position (1) but not at any of the surround positions (2–9) was taken to be the extent of the classical receptive field of the cell. The texture was sufficiently dense so that the receptive field of the cell was covered by multiple line segments. The patch stimulus displayed is 1.5° (30 pixels) in diameter and contains multiple short line segments. An additional method to assess the spatial extent of the RF was to test the cell with a hole stimulus of different diameters, centered on the receptive field. The diameter at which the cell's response dropped precipitously was considered the maximum extent of the receptive field. One cell's response to hole and disk stimuli is shown in the illustration. The RF extent of this cell measured moving bar method was 0.75°, and by the patch method was 1.0° in diameter. Note that the response to the disk and hole stimulus was significantly decreased at 2.0° diameter.

ron's receptive field is encoding a part of the figure or a part of the background.

In the following two sets of experiments, we studied this figure-ground hypothesis further by examining the precise nature of the interior enhancement and the conditions that gave rise to this phenomenon. We also examined the responses of the neurons at different time windows to elucidate the spatiotemporal dynamics of the neurons under various testing conditions to different stimuli, and looked for neural activities that might reflect higher-order perceptual computations in V1.

3.2. Methods and materials

In this series of experiments, we recorded from 301 neurons in the parafoveal area V1 of three awake behaving rhesus monkeys *Macaca mulatta* while the monkeys were doing the following fixation task. At the beginning of each trial, a monkey first established fixation to a red dot on the screen within a 0.3° fixation tolerance window for 200 ms. Then, a stimulus was flashed on the grey screen for 350 ms as the test stimulus. The monkey was kept alert by being required to make a saccadic eye movement to a target that appeared in a random position upon the disappearance of the test stimulus and the fixation dot. Correct saccades were rewarded with drops of apple juice. Eye movements were recorded using implanted scleral search coils [31] and sampled at a rate of 200 Hz.

3.3. Stimulus display and recording methods

Stimuli were presented on an NEC multisync XL color video display monitor, driven by a Number Nine Corporation SGT Pepper graphics board with a 640 × 480 pixel resolution, at a frame rate of 60 Hz. The screen was 32 × 24 cm in dimensions and was viewed from a distance of 58 cm. One pixel thus corresponded to a visual angle of 0.05°, and the full screen size was 32 × 24°.

Recordings were made transdurally with glass coated platinum-iridium micro-electrodes through a surgically implanted well overlying the operculum of area V1 of awake behaving monkeys. The recording well was surgically implanted when the monkeys had acquired a sufficient level of performance. All surgical procedure were performed under deep pentobarbital anesthesia and all experimental procedures were in accordance with NIH guidelines (see also ref. [32]). Impedance of the electrodes ranged from 1.0 to 4.0 MΩ. Spikes from single units or in some cases clusters of several units were isolated either by setting an amplitude threshold or by cluster cutting using the DataWave system.

The classical receptive field of each neuron was localized, its orientation preference and spatial extent ascertained by slowly moving a black thin bar on the screen in different directions over the receptive field. The maximum spatial extent of the classical receptive field was further assessed by successively flashing circular texture patches at and around the localized receptive field until only the center patch would elicit response in the neuron (Fig. 2). An additional method was to project a hole stimulus onto the receptive field. The diameter of the hole was increased in successive trials until no significant initial outburst of response was elicited in the neuron. The receptive fields mapped by the bar were slightly smaller than those mapped by the texture patches. Generally, we take the classical receptive field to mean the area of the region where direct stimulation by bar or texture will produce a strong and

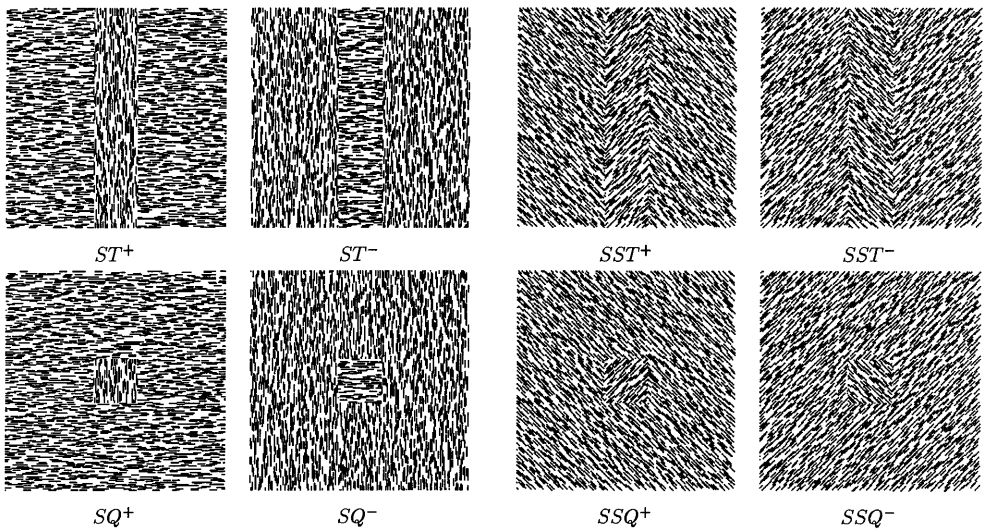

Fig. 3. Four pairs of complementary stimuli are shown in the display: texture strip (ST^+/ST^-), slanted texture strip pair (SST^+/SST^-), square pair (SQ^+/SQ^-). slanted texture square pair (SQT^+/SQT^-). In the positive (+) stimulus, the figure contains texture of the preferred orientation of the cell being tested. In the negative stimulus, the figure contains texture of the orthogonal (anti-preferred) orientation.

brisk response, while stimulation of its surround alone (e.g. patches 2–9) will produce a negligible response.

3.4. Experiment I: neuronal response within texture figures

3.4.1. Motivation

Lamme's figure-ground hypothesis [25] was based on several observations. First, the neurons' responses were enhanced uniformly within the square texture figure, i.e. the enhancement at the boundaries and at the interior of the figures were more or less the same. Hence there was a sharp asymmetry in the enhancement at the figural border between the figure and the background. The later stage of V1 responses correlates primarily with the figural signal and was independent of the receptive field size and orientation preference of the cells. The fact that the enhancement could be induced either by texture cues or motion cues further suggests the enhanced neural activities might be used to represent a more abstract perceptual structure.

This evidence is not completely consistent with Gallant et al.'s [33] findings that V1 neurons were sensitive to texture contrast edges. Furthermore, the evidence that the enhancement at the later response is insensitive to the orientation tuning of the cells is also counter-intuitive since one would expect that orientation-selectivity of V1 neurons would continue to play an important role in higher order contour completion at the later stage of their responses, for example, in the task of contour completion. One potential problem with Lamme's [25] original experiment was that a fairly large (1°) fixation tolerance window was used. Could the uniform interior enhancement observed arise from the smoothing-in of the edge signals? Therefore, we conducted the following experiments using a much smaller fixation window (0.3°) to elucidate the relationship between edge enhancement and the interior 'figural' enhancement, and the role of the orientation-selectivity of the cells in mediating all these effects.

3.4.2. Methods

To elucidate the relationship between edge enhancement and the interior enhancement, we tested the responses of V1 neurons to three main sets of stimuli: texture boundary stimuli (Fig. 5), texture strip stimuli (ST, SST), and texture square stimuli (SQ, SSQ) (Fig. 3). These stimuli were tested in complementary pairs as shown. The width of the strips and the squares was 4° visual angle, which was about four to six times the size of the cells' receptive fields at parafoveal eccentricity of 3–4°. At a later stage of the experiment, neurons were also tested with strips of different widths, and with texture shapes such as diamonds and rectangles for reasons to be described later. The neurons were tested under one or more of the following three testing conditions: parallel, orthogonal and oblique, which specified the relative difference in orientation between a cell's preferred orientation and the orientation of the figure boundary it encountered (Fig. 4). To maximize our study of the cells in the parallel condition, we frequently rotated the strips so that the boundary became parallel to the preferred orientation of the cells.

The location of the fixation spot, and hence that of the receptive field of each cell, was kept constant across trials. The figure was presented at different translated positions relative to the classical receptive field in successive trials. The responses of the neurons when their receptive fields were located at the boundary, interior and exterior of the figure were studied in successive

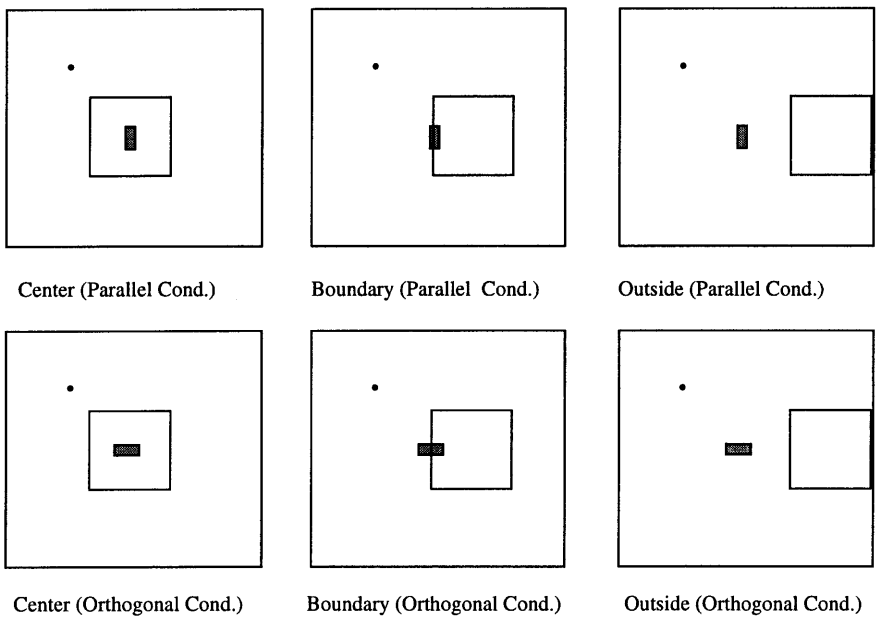

Fig. 4. The displays illustrate the placement of the square figure relative to the receptive field and the preferred orientation of a cell under the parallel and the orthogonal testing conditions. Each frame depicts a particular stimulus configuration on the monitor relating the fixation spot (the black dot), and the cell's oriented receptive field (the gray rectangle) to the figure in the stimulus (the square). Here, the square figure is shown to be displaced horizontally over successive trials so that the cell's receptive field was placed at the center, the boundary, and outside of the figure in different trials. In the parallel condition, the preferred orientation of the cell was parallel to the figure boundary it encountered. In the orthogonal condition, the preferred orientation of the cell was orthogonal to the figure boundary it encountered. The sampling line is defined as the line on which the receptive field of the cell is translated over trials. In these diagrams, it is horizontal. In order to study the neurons' responses in the parallel condition, we frequently rotated the strips and the sampling line so that the strip was parallel to the preferred orientation of the cell and the sampling line was orthogonal to it. In successive trials, nine evenly spaced positions within the figure of each stimulus and seven positions outside the figure along a sampling line were presented to the cell at a spatial interval equal to 1/8 of the width of the figure.

trials (Fig. 4). In order to reveal the neurons' sensitivity to image structures other than their orientation-selectivity to local features (orientation tuning), the responses of each neuron to a stimulus (e.g. SQ^+) and its complement (e.g. SQ^-) were summed at each corresponding position to produce a combined response that is independent of the orientation tuning of the cell (illustrated in Fig. 7, also see ref. [25]). Each stimulus pair was tested within each block of the experiment. The two stimuli in each pair and sampling positions were randomly interleaved during the presentation. In these experiments, the exact texture pattern projected onto the receptive field was kept constant so that the variability in the response of the neuron was not due to the precise nature of texture patterns but to the contextual effects. This requirement however could not be kept at the texture contrast boundary.

3.4.3. Results

The data presented here were drawn from 219 neurons. These neurons were primarily complex cells, sensitive to luminance contrast of both polarities. We found that there were several stages in the responses of V1 neurons, with distinct spatial response profiles at different temporal windows. Typically, V1 neurons started to respond about 40 ms after the stimulus was displayed on the screen. From 40 to 60 ms after stimulus onset, the cells behaved essentially as local feature detectors or linear filters [9,34]. The responses to the texture stimuli were therefore initially uniform within a region of homogeneous texture based on the orientation tuning of the cells. At 60 ms after stimulus onset, boundary signals started to develop at the texture boundaries. By 80 ms, the responses at texture boundaries have become sharpened (Fig. 5), consistent with the psychophysical time course of texture segmentation [35].

Fig. 6 shows examples of the orientation-specific responses of two neurons to the texture strip pair ST^+ and ST^- in different time windows. From 40 to 80 ms, the neurons responded uniformly well within the interior of the positive strip (ST^+), but responded very poorly outside the strip. This was because the texture within the strip was of the preferred orientation of the cells, and the texture outside was of the non-preferred orientation. The situation was reversed in the negative strip (ST^-) in which the neuron was found to respond primarily to the preferred texture outside the strip. Interestingly, 80 ms onward, as the responses at the boundary became more localized, a response peak was sometimes observed at the center or the axis of the strip when the strip stimuli (ST^+ and ST^-) were tested

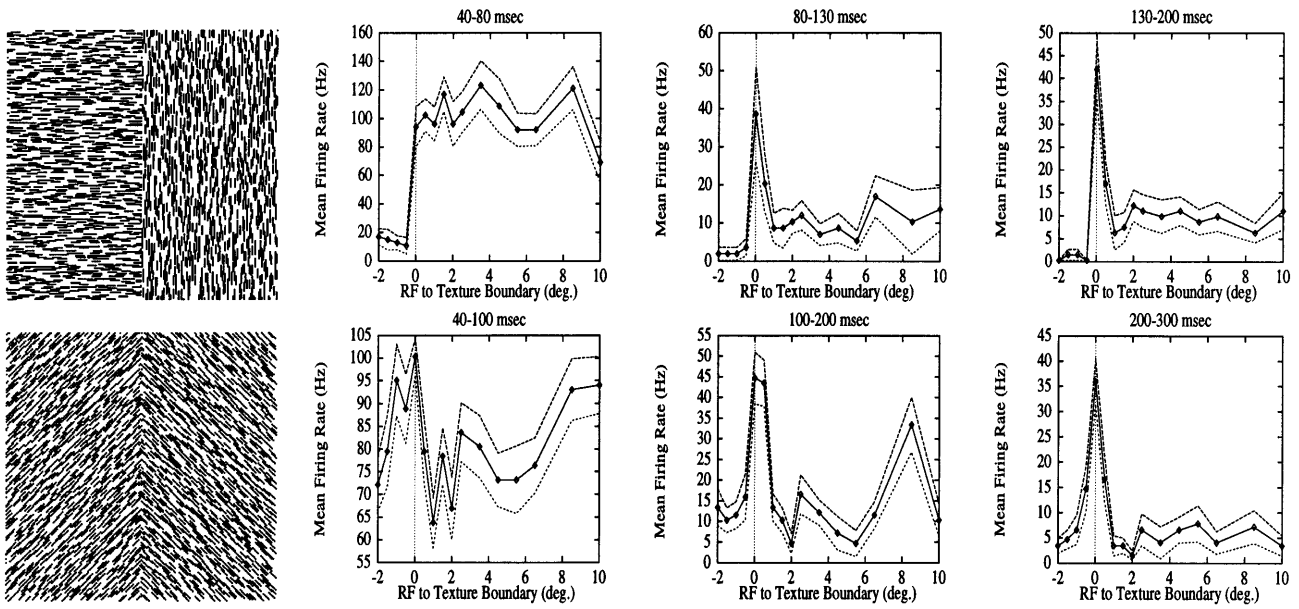

Fig. 5. Top row: the spatial response profile of a V1 neuron to the texture boundary stimulus at different time windows. The preferred orientation of the cell was vertical. The texture in the region to the right of the boundary is of the preferred orientation of the cell. The solid lines in these graphs indicate the mean firing rate within the specified time window, and the dashed lines depict the envelopes of S.E. The dots on the solid lines are the data points. The abscissa is the distance in degrees of visual angle from the RF center to the texture boundary. Bottom row: the spatial response profile of another vertical cell to a texture boundary defined by slanted texture. six out of ten neurons tested with such texture boundary showed similar sharpened responses, suggesting that some V1 neurons are sensitive to texture boundaries regardless of the orientation of the defining texture.

under the parallel condition. The spatial response profiles in successive temporal windows show the development of these central peaks (Fig. 6). In one dramatic example, cell m32 did not respond at all within the strip in stimulus ST^- in the initial phrase but became active at the axis of the strip after 80 ms. Statistically significant central peaks were observed in 14 out of the 50 neurons tested with strip ST^+ and 10 cells with strip ST^- (T-test, $P < 0.05$).

Consistent with Lamme's [25] observation, we found that starting at around 80 ms, the combined neural responses were higher when the receptive fields of the cells were inside the square in stimuli SQ^+ and SQ^- than when they were outside the square. However, we found that particularly in the parallel condition, the differential between the responses inside and outside of the figures was not uniform, but layered or characterized with additional structures. First, there were the boundary enhancement signals, which were about four times stronger on the average than the interior enhancement signals in the parallel condition (Figs. 8 and 9). Second, we again observed an extra response at the center of the square in the parallel condition and sometimes in the slant condition. Fig. 7 illustrates the responses of the individual cell to both square and strip, showing the center peak response in both cases. The response within the square figure was significantly higher than the response within the strip figure. The combined response, obtained by summing the response to positive and negative figures (Fig. 7), showed that the neuron experienced much stronger interior enhancement within the square than within the strip. Forty five cells tested with square (SQ^+/SQ^-) in the parallel condition showed the same sharp pronounced boundary responses as the cells' response to (ST^+/ST^-). While only eight of 50 cells showed minor overall interior enhancement within strip (ST^+/ST^-), 32 out of the 45 cells tested with squares showed statistically significant interior enhancement (T test, $P < 0.05$). Statistically significant central peaks were observed in eight neurons for SQ^+ and nine for SQ^-. The population histograms of these effects are shown in Fig. 9.

The temporal response of the different groups of cells, classified according to orientation selectivity and testing conditions, at several selected positions of the strip figures were shown in Fig. 8 to illustrate the boundary response, texture surface response, inside enhancement response and the axis response. Fig. 9 shows the histograms of the cell distribution for these effects. We observed a slight inside-versus-outside enhancement for the strip stimuli in the non-parallel (orthogonal and slant) testing conditions, but not much at all in the parallel testing condition, whereas the inside-versus-outside enhancement was observed in squares in all conditions.

Fig. 10 shows a series of 3D graphs depicting the spatiotemporal profiles of the combined responses of V1 neurons to different stimuli under different condi-

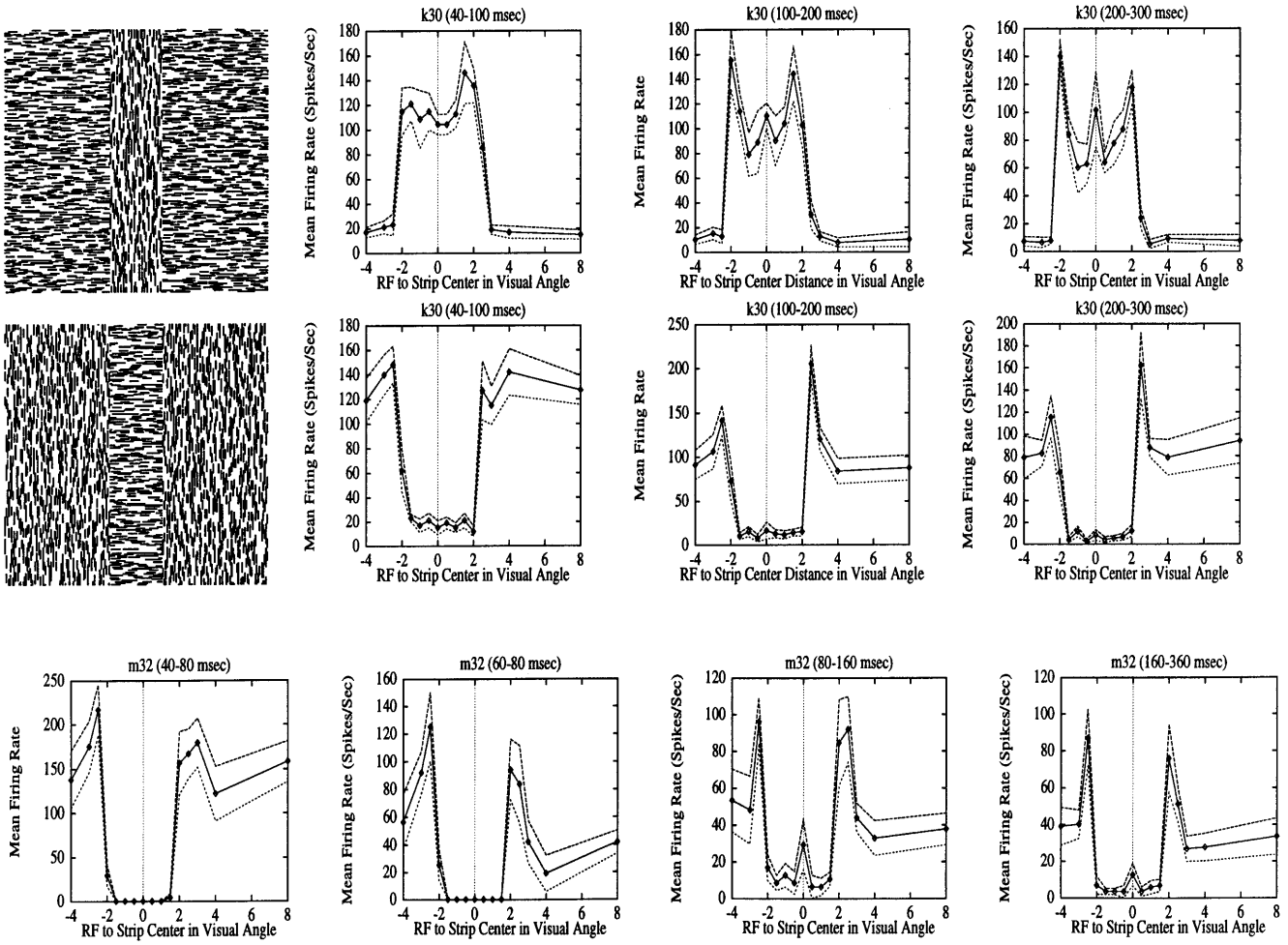

Fig. 6. Shown here are the spatial response profiles of two vertical orientation selective V1 neurons to different parts of the strips (ST^+ and ST^-) along a horizontal cross-section (sampling line) in different time windows after stimulus onset. The width of the figure was 4°. The receptive fields were at about 5° eccentricity in the visual field, and about 1.2° in diameter. The abscissa is the distance in visual angle from the RF center to the center of the figure. The solid lines in these graphs indicate the mean firing rate within the time window, and the dashed lines depict the envelopes of S.E. The strips were defined by texture contrast. Response of cell k30 to both positive and negative stimuli (see Fig. 3 caption) are shown. The spatial response profiles of another V1 neuron (m32) to the negative strip ST^- stimulus at different time intervals are shown in the bottom row. Approximately 40–60 ms after stimulus onset, the cell responded uniformly to the background and did not respond to the texture strip at all because it was not tuned to the texture inside. From 60 to 80 ms, the boundary started to sharpen, but there was still no response within the strip. Interestingly, 80 ms onward, a pronounced response peak gradually developed at the axis of the strip.

tions. These graphs revealed some additional detailed information that is difficult to detect from the responses depicted in 2D time windows. The combined initial responses to the texture stimuli during the initial 40–80 ms period were shown to be uniformly colored across space, corresponding to the orientation specific response to local features. After the initial burst of response, there was a transient drop in neural activity, followed by a resurgence at about 80–100 ms. It was during and after this period of resurgence that we observed the several remarkable phenomena in the population that might be related to higher order perceptual computations: i.e. the cue-invariant boundary responses (Fig. 10A and B), the extra responses at the center of the strip (Fig. 10C), and the boundary and inside-versus-outside enhancement of the neurons within the square under different testing conditions (Fig. 10D,E and F). The center response was most evident in some cells for the strip and the square in the parallel condition. However, it was also evident, discontinuously, in the population response average of neurons to slant texture square (SSQ^+/SSQ^-) in the slant testing condition (Fig. 10E). Another remarkable observation is that when the vertical cells were tested with slanted texture strips SST^+ and SST^- in the parallel condition, the cells' initial responses were relatively mild because both the textures inside and outside the figure were not of the preferred orientation of the cells. Remarkably, texture boundary signals that emerged during the resurgent period were actually stronger than their initial responses (Fig. 10B), suggesting that the later response of the cells were more specific to the

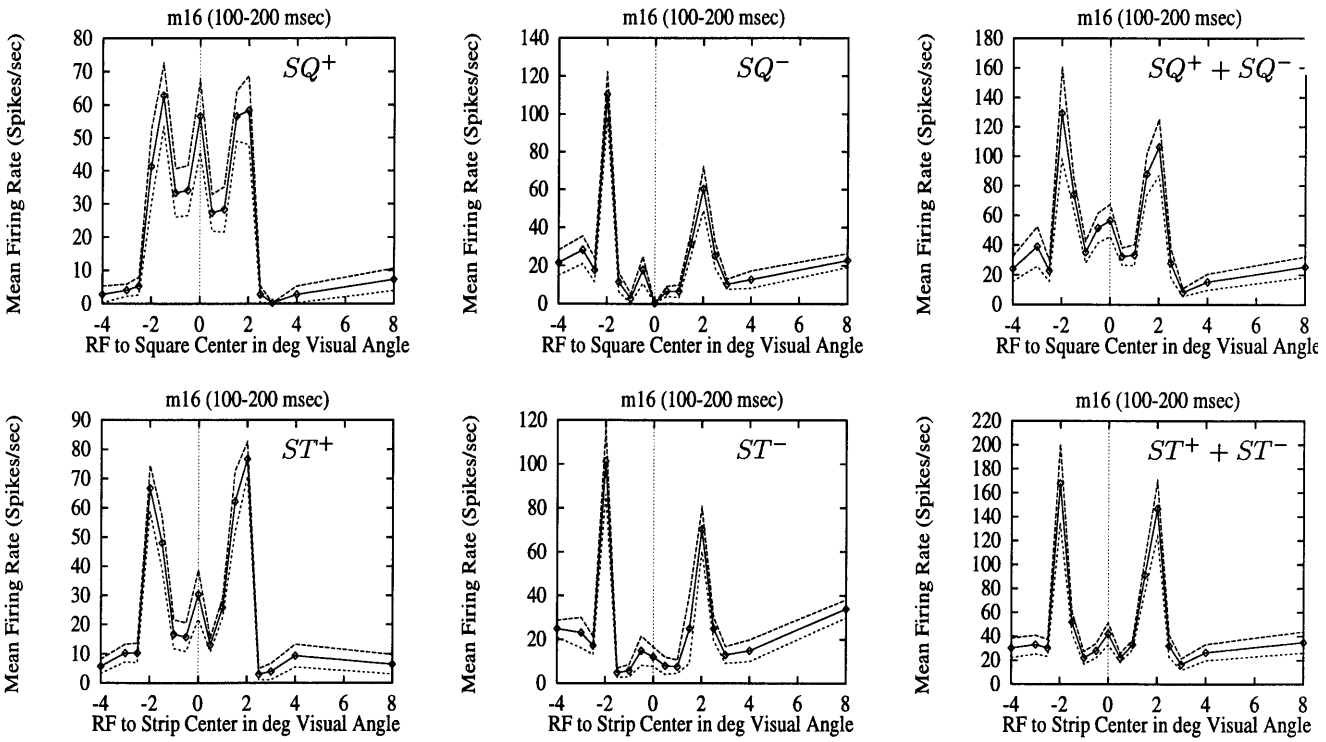

Fig. 7. Spatial response profile of a V1 neuron (hl6) 100–250 ms after stimulus onset responding to the strip as well as the square stimuli of 4° width. Boundary responses were sharply localized at the boundary positions. The center peaks were observed at the center of both ST^+ and SQ^+, but slightly shifted in ST^- and SQ^-. The response within the strip of ST^+ is significantly less sustained than that within the strip of SQ^+. The third column illustrates how signals from the complementary stimulus pair are combined: the response to the positive stimulus (e.g. SQ^+) and to the negative stimulus (e.g. SQ^-) were summed at each spatial location. The resulting combined response for the square shows substantially greater neuronal response inside the figure than outside the figure (inside/outside enhancement). The inside/outside enhancement was not significant in the strip. Both showed a response peak at the center for the positive figures, but shifted from the center by 0.5° in the negative figures.

orientation of the boundary than to the local texture features.

The response enhancement at the axis were significant in 25–30% of the neurons. But the effect seems to be less significant in the population response average shown in Fig. 10A, though a hint of the axis response was evident. The magnitude of the axis responses were comparable to the interior responses, and were significantly less than the boundary responses. When strips of different widths were tested, we found that most 'center-responding' neurons exhibited central peak responses only for a narrow range of widths. The central response peak tended to appear at a particular strip width for an individual neuron and disappeared as the strip became wider or narrow. However, we found, for each strip width, there were neurons producing the central response peaks. Examples of neurons responding to the center of 2, 3, 4 and 6° strips were shown in Fig. 11.

It is important to examine the spatial distribution of the response peaks inside the figure, for the null hypothesis is that the response peaks were simply randomly distributed. The metric used in Fig. 9 serves only to show the relative magnitude of a central response peak, but cannot prove the response peaks were always centered. In fact, some response peaks were shifted from the center (see the response of a cell to the negative strip and square in Fig. 7). To address this concern, we plotted the histogram of the spatial locations of statistically significant interior response peaks of the neurons studied (Fig. 12). The histogram shows that there was a certain degree of dispersion of the interior response peaks from the center (slightly stronger than the dispersion of the boundary locations), but there was a strong emphasis on the center. Curiously, the histogram also showed an increased separation in the boundary responses of the negative strip, which might be related to the perceptual widening effects of horizontal stripes well known to fashion designers.

When more complex shapes such as diamond, rectangle, and overlapping rectangles were tested, central response peaks could be detected in 10% of the cells in diamond, and 30% of the cells in rectangle and overlapping rectangles. The center-responses were particularly pronounced at the center of mass of the diamond and of the rectangle (Fig. 13). The axis-responses of the neurons in the case of overlapping rectangles suggest

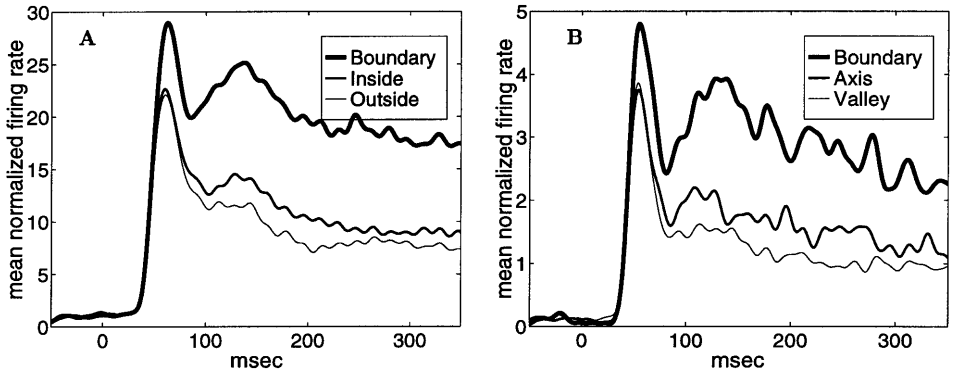

Fig. 8. (A) The averaged temporal response profile of the 50 cells responding to square stimuli SQ^+ and SQ^- in the parallel condition, The width of the square was 4° visual angle. The average RF size of the cells was about 1°. The boundary signal was the average of the responses at two positions -2 and 2°, the inside signal was the average of the responses between positions -1 and 1° within the square. The outside signal was the average of the responses at positions -4, 4, 8°. Each neuron's response to the two stimuli in each complementary pair was combined at each point in time and space and then normalized by the mean of the cell's maximum response at all positions across time. The signals were then averaged across the population of neurons and smoothed with a Gaussian ($\sigma = 8$ ms) in time. There was no smoothing in space. This graph shows that the boundary enhancement signal was about four times greater than the interior enhancement signal in the parallel condition. (B) The averaged-normalized temporal response profile of the 14 neurons that exhibited response peaks at the center of the strip ST^+, showing a significant enhancement at the axis relative to the adjacent positions inside the strip. The axis signal was the response at position 0°, and the valley signal was given by $[\min(R_{0.5}, R_1) + \min(R_{-0.5}, R_{-1})]/2$, where R is the response at the location indicated in the subscript. The temporal response profile shows that the differential response between the axis and the valley points within the strip emerged at about the same time the inside signals became differentiated from the outside signals.

that central responses might be computed after the figure-ground relationship has been determined (Fig. 13).

3.4.4. Discussion

There are two main new findings in these experiments. First, we showed that when the neurons' preferred orientation was parallel to the figure boundary, significant enhancement was observed at the boundary regardless of the nature of the texture contrast (Fig. 10). The boundary enhancement was significantly stronger than the interior enhancement (Figs. 8 and 9). This shows that the orientation-selectivity of the neurons is important even at the later stages of V1 responses. When the neurons' preferred orientation was orthogonal to the boundary, the difference between the boundary enhancement and the interior enhancement was less dramatic, suggesting that the coding of surface qualities was emphasized in such a condition. These results suggest that V1 neurons are signaling at least two types of information in their spike trains: boundary location and surface qualities.

Second, we confirmed Lamme's [25] observation that there is significant interior enhancement within the figure but found additional spatial structures in the interior enhancement. The interior enhancement occurs regardless of the orientation of the cells, and is relatively spatially uniform within the figure. However, when the orientation preference of the cell is parallel to that of the figure boundary as well as the texture inside the figure, we observed additional significant enhancement of responses at the boundary as well as at the center/axis of the figure in about 30% of the cells. The enhancement within the figure and the extra enhancement at the center or axis of the figure resonate with Kovacs and Julesz's [36] psychological findings that human perceptual sensitivity was enhanced within compact figures and was markedly enhanced at the center of a circle. Kovacs and Julesz [36] have suggested that the central enhancement might be related to the medial axis transform. Therefore, the enhancement of the neurons within the figure might have perceptual and computational significance, perhaps signaling the closure of contours around the surface of a figure, and possibly the location of the axis of symmetry of an object.

The medial axis transform is a robust method for describing a local region. By definition, medial axis is the locus of the centers of the largest disks inside a region (Fig. 14). Hence, medial axis transform contains two pieces of information: the diameters of the disks (scale) and the locus of the disks' centers (location). Using this method to describe a region, a neuron needs to be sensitive to the conjunction of three relatively global features: there has to be at least two distinct boundary segments touching a disk of a certain radius, and the disk interior has to be homogeneous in surface quality (Fig. 14).

There are several ways to compute and represent the medial axis transform. One is the grassfire algorithm proposed by Blum, which mimics a grass fire that propagates in from the boundary of a figure; the points of collision form the skeleton of the figure, and the fire itself colors the figure. The magnitude of the response at the location of the axis might encode the diameter of

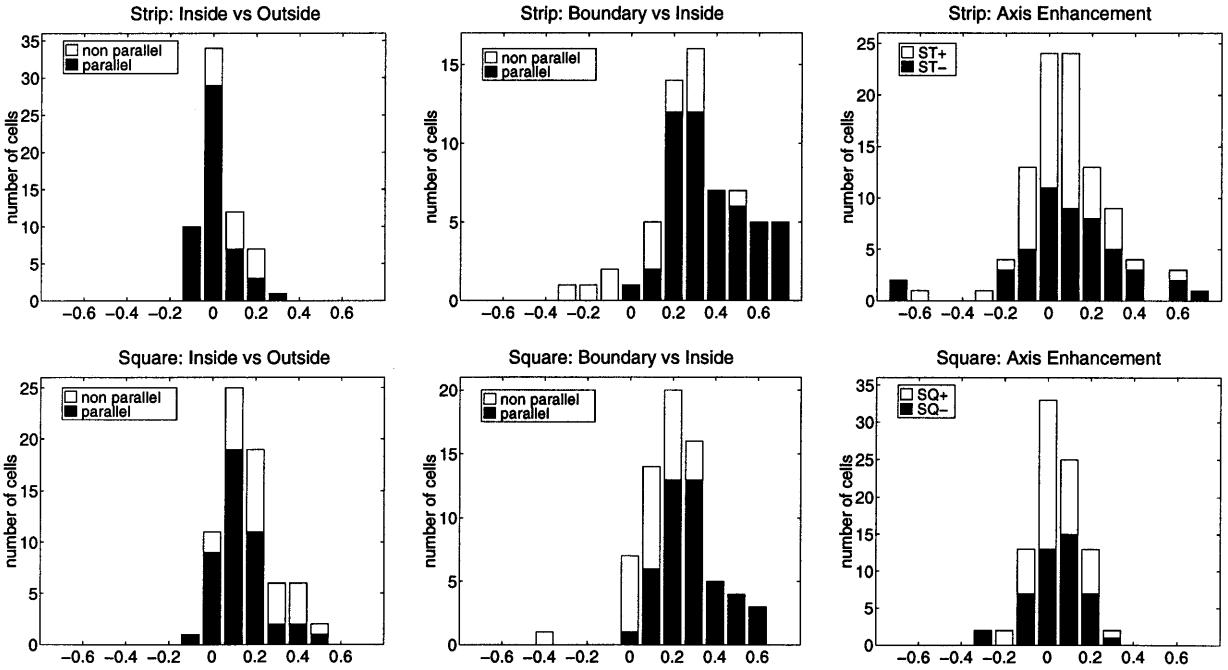

Fig. 9. Population histograms of the cells illustrating the prevalence of the various effects. All the modulation ratios, shown on the abscissa, were computed by $(A-B)/(A+B)$. For the inside versus outside modulation, A was the average of the responses inside the strips or squares, and B was the average of the responses outside the strips or squares (see Fig. 8A caption). The histograms show that the interior enhancement effect was strong for the squares but weak for the strips. For the boundary versus inside modulation, A was the average of the responses at the two boundary positions, and B was the average of the interior responses (see Fig. 8A caption). The histograms show significant boundary responses over interior responses for both the square and strip under the parallel condition. The cells in the orthogonal (nonparallel) condition emphasized on the surface aspects of the signals, while in the parallel condition they emphasized the boundary signals, but carried other aspects of the signals as well. For the axis enhancement, A was the average of the responses at the center of SQ^+ or SQ^- or the axis of ST^+ or ST^-, and B was the average of the responses of the valley points surrounding the central peaks (see Fig. 8B caption). The histogram shows that, particularly for the strip stimulus, the neural response to the center was enhanced relative to the responses at the positions adjacent to the center.

the disk or the distance away from the borders. Alternatively, Crowley and Parker [37] envisioned a center-surround mechanism, similar to a Laplacian of Gaussian filter of a large spatial extent, to compute the skeletal ridges in images. Burbeck and Pizer [38] proposed a similar but distinct mechanism in which a V2 neuron extended its two dendritic 'arms' to gather border signals from V1 neurons on its two sides. In these later models, the medial axis was thought to be computed and represented by an ensemble of neurons, each 'tuning' to a portion of the region of a particular size or width. Our data show that individual neurons exhibit central response peaks primarily for a particular width, but not along the entire medial axis (e.g. diamond). This evidence suggests that the location information about the medial axis, if it exists, is likely encoded by an ensemble of neurons in V1. The following experiment investigates the medial axis hypothesis further by examining the dependence of central enhancement and location of the response peak on the size and the defining cue of the figure.

3.5. Experiment II: neuronal responses within uniformly colored figures

3.5.1. Motivation

Experiment I shows that the neural response within a texture square are much stronger than that inside a texture strip. Therefore, V1 neurons were obviously sensitive to the area of a local texture region. However, the response peak was sensitive to the width of the region (or the diameter of the largest disk) but less critically to the area since the central response peaks could be observed for both the square and the strip stimuli in an individual neuron (Fig. 7). However, it was not clear whether the sensitivity to the region's spatial extent (area) and the central peak responses were unique to texture figures, or were more general and abstract. Although there was evidence that the 'color blob' cells [39] are capable of responding to colored disks, the spatial extent of their sensitivity was not fully characterized. The objective of this experiment was to understand the relationships between the interior enhancement of the neural responses, the size of the figure, foreground and background texture, and the

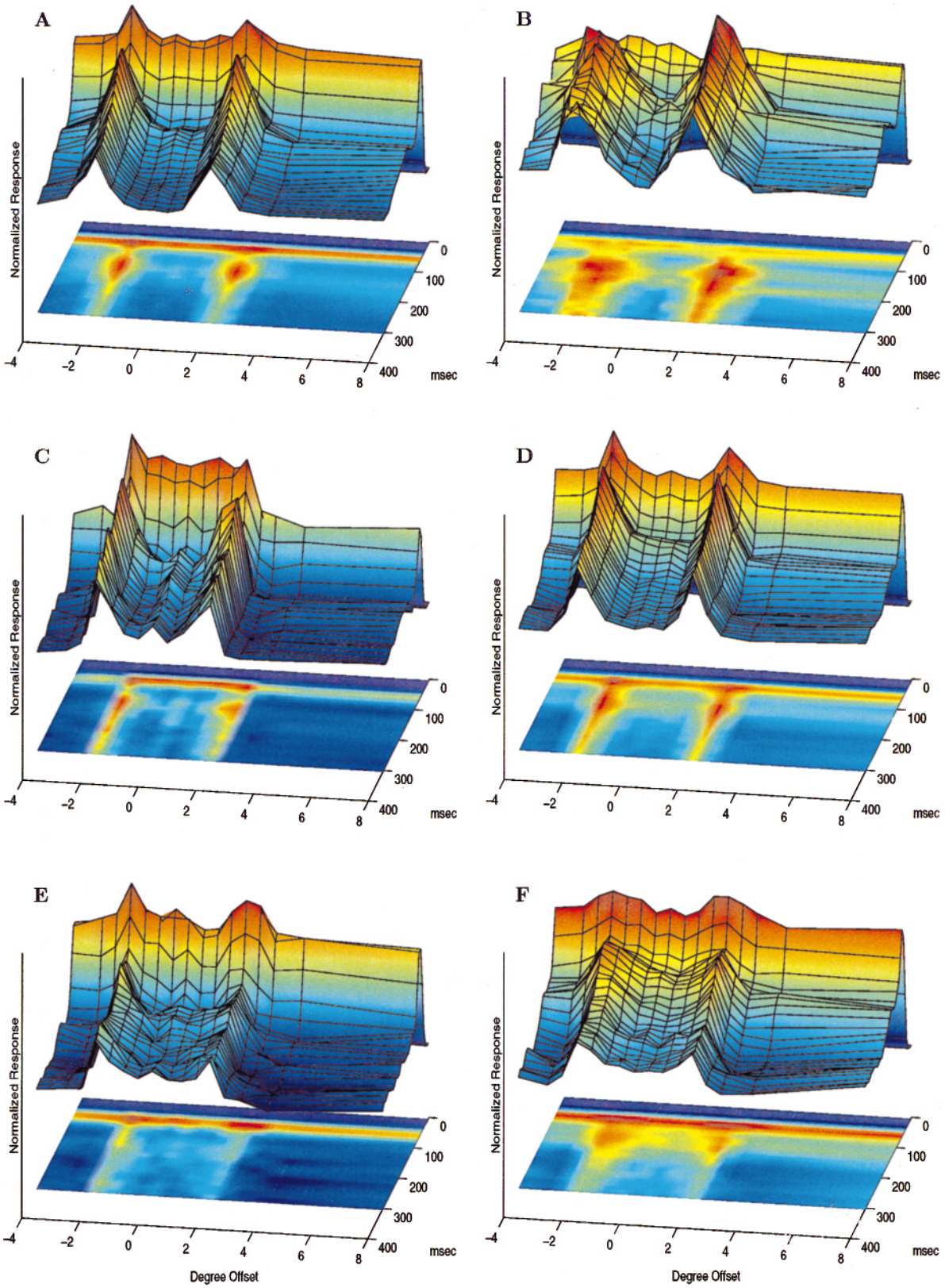
A
B
C
D
E
F
Normalized Response
0
100
200
300
400
msec
−4
−2
0
2
4
6
8
Degree Offset

Fig. 10.

cues that made up the figure. In particular, we would like to know whether the interior enhancement and the central peak are observed in neurons responding to uniformly colored figures.

3.6. Methods

The stimuli used in this experiment are depicted in Fig. 15. They were gray, white and black disks of various sizes, ranging from 29–10° visual angles in diameter. The grey disks were surrounded by textures that were parallel or orthogonal to the preferred orientations of the cells. Texture disks were also tested for comparison. The disks were surrounded either by contrasting textures of 45 or 90° orientation differences, or by gray background. The receptive fields of the cells in one monkey ranged from 1–1.5° in size at 5–6° eccentricity and in another monkey ranged from 0.5–1° in size at 2.5–3.5° eccentricity in the visual field. The center of the displayed disk was placed on top of the classical receptive field. The minimum size of the disk (2° diameter) was significantly larger than the typical receptive field size (1–1.5°) in this parafoveal region of V1. The screen during the inter-trial interval was gray. Therefore, the cells at the center of the white or black disks experienced a step change in luminance at stimulus onset. The cells at the center of the gray disk experienced absolutely no change in visual stimulation within the receptive fields at all.

The spatiotemporal response profile of neurons to black/white strips were obtained using the procedure of Experiment I to investigate whether central peaks could also be observed in uniformly colored strips (Fig. 19). A small number of neurons were also tested with a set of stimuli shown in Fig. 20.

3.7. Results

A total of 83 neurons from two monkeys were studied in this set of experiments. Forty two neurons were tested with black, white, gray and texture contrast disks of three scales: 2, 3.7 and 7° in diameter. Twenty neurons were tested with gray disk and texture contrast disks of finer scale, ranging from 1 to 9.5°. Twelve neurons were tested with black/white strips of 3° width (Fig. 19). Nine neurons were tested with the ring and texture contrast disk stimuli (Fig. 20).

Many of these cells responded to the center of uniformly colored figures (Fig. 16). Because there were no oriented features within the classical receptive fields to stimulate the cells, the firing rates of the cells were naturally very low (typically 5–15 spikes/s compared with 60–120 spikes/s for texture disks). However, the response inside a uniformly colored figure is significantly greater than the corresponding uniform background, exhibiting similar size-dependent interior enhancement as the texture disk (Fig. 17A). The inside/outside differential responses emerged around 60–70 ms for the uniformly colored disks, slightly earlier than the texture disks (Fig. 16). Ten percent of the cells exhibited interior enhancement for both texture defined disks and uniformly colored disks (Fig. 17B). Fig. 17(A) shows that the interior enhancement was actually stronger for the uniformly colored disks than for the texture figures, perhaps because of their stronger perceptual saliency. As the size of the figure increased, the responses at the center of the figures decreased for both textured or uniformly colored disks. In general, the interior enhancement was a function of both figural size and the eccentricity of the neurons. A 4° wide square can induce less than the average.

When the neurons' responses to the gray disks and texture disks were studied with finer size resolution, two additional observations could be made. First, while the neurons were sensitive to the orientation of the texture within the texture figure, their responses were insensitive to the orientation of texture outside a texture or gray figure (Fig. 18). Second, while the boundary signals within the texture signals 'contracted' toward the location of the borders over time, the boundary signals of the gray disks spread inward from the borders over time (Fig. 18).

The spatiotemporal responses of the 12 neurons tested with black and white strips showed the formation of

Fig. 10. Spatiotemporal response profiles illustrating the dynamics of the V1 neurons to the stimuli under the different conditions. The signals in most of these 3D plots were obtained by the procedure described in Fig. 8(A) caption. Fig. 10(C) shows the population response of axis-positive cells to stimulus ST^+ only. (A) The responses of 50 cells to the texture strips ST^+/ST^- in the parallel condition showed a uniform initial response across space, followed by the emergence of localized boundary signals at the texture boundaries. A hint of axis enhancement is evident between 80–120 ms even in the average of the entire population. This phenomenon was even more striking in the selected cells. (B) The responses of 10 cells in the parallel condition to slanted texture strips SST^+/SST^- showed an initial phase of mild response was followed by a stronger boundary response after 100 ms. (C) The average response of the 14 selected neurons, out of the 50 neurons tested with texture strip, that showed statistically significant central response peaks to strip ST^+ in the parallel condition. This was characterized by a persistent central ridge throughout the duration of the later response in their spatiotemporal response profile. (D) The responses of 45 neurons to the square SQ^+/SQ^- in the parallel condition showed that interior enhancement became noticeable at 80 ms, together with the significant boundary responses which were sharply localized, pronounced and persistent throughout the duration of the trials. (E) The responses of 15 cells in the slant condition to the slant texture squares SSQ^+/SSQ^-. The texture inside and outside the square was parallel or orthogonal to the preferred orientation of the neurons, and was slanted (45 or 135°) from the orientation of the boundary (see SST^+/SST^- for example). Significant inside/outside enhancement was observed. Interestingly, a small central response peak was observed on and off during the evolution of the average response of the whole unselected population. (F) Sixteen cells tested with squares SQ^+/SQ^- in the orthogonal condition revealed that the later part of their responses was characterized by a more uniform interior enhancement response within the figure than that in the parallel condition.

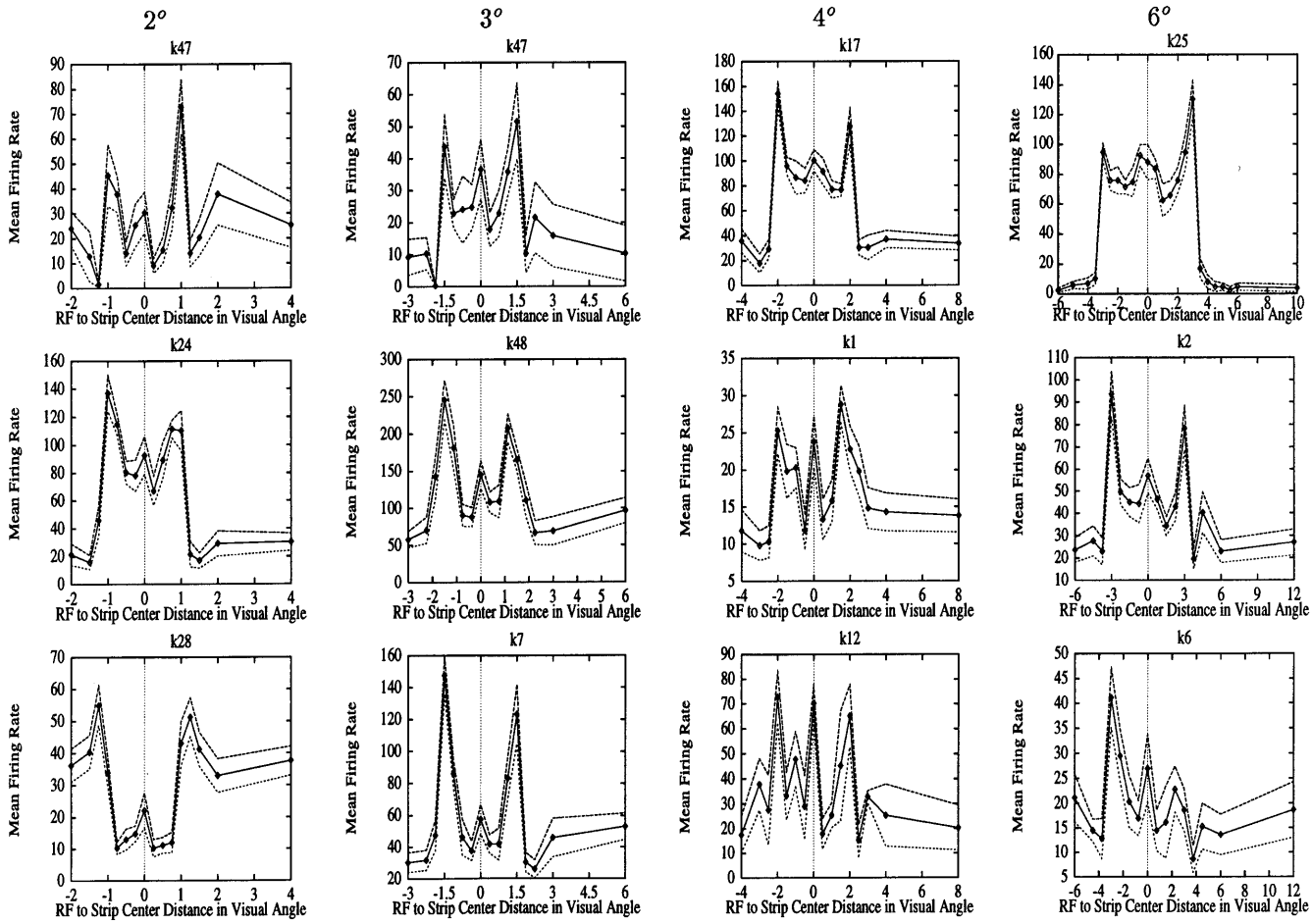

Fig. 11. The display shows examples of neurons exhibiting central peaks at different strip widths. Statistically significant central response peaks were observed in either positive or negative strips in three out of 11 cells tested with a 2° strip, four out of 12 cells tested with a 3° strip, 14 out of 50 cells tested with a 4° strip, and five out of ten cells tested with a 6° strip. Cell k47, as shown, exhibited central peaks for both 2 and 3° strips.

luminance border signals almost instantaneously at the beginning of the response onset, i.e. 40 ms after stimulus onset. Responses within the black/white strips showed slight interior enhancement, but the responses in general were very weak. No statistically significant central response peaks could be observed (Fig. 19).

The responses of nine neurons to stimuli in Fig. 20 were remarkable. When a texture disk was surrounded by black background, the responses of the neurons at the center of the disk were markedly enhanced compared to their responses at the center of the disks defined by texture contrast. However, when the texture border of the texture contrast disk was occluded by a black ring, the responses were markedly suppressed (Fig. 20).

3.8. Discussion

This set of experiments provide several new insights. First, the interior enhancement is a general phenomenon. It was observed for uniformly colored figures, and was not limited to texture figures. Second, the luminance/color or texture contrast borders of the uniformly colored figures were capable of inducing the interior enhancement signals in the receptive fields from afar. The responses of the neurons were sensitive to the size of the figure regardless of whether the figures were uniformly colored or textured. Third, the responses of the cells to a uniformly colored figure were insensitive to the orientation of the texture in the surround. These facts suggest that the interior enhancement signals arise from the more abstract border signals of the figure, not merely a consequence of lateral inhibition or excitation by texture elements inside or outside the figure.

The evidence shows signals sharpen spatially over time at the boundaries of the texture figures, and spread spatially over time from the boundaries of uniformly colored figures. This suggests that there are at least two underlying processes mediating these effects: a lateral inhibition process for texture boundary detection, and a lateral excitation process, possibly for surface interpolation.

Although interior enhancement could be detected in uniformly colored figures, the responses were too weak to discern any spatial structure such as the central

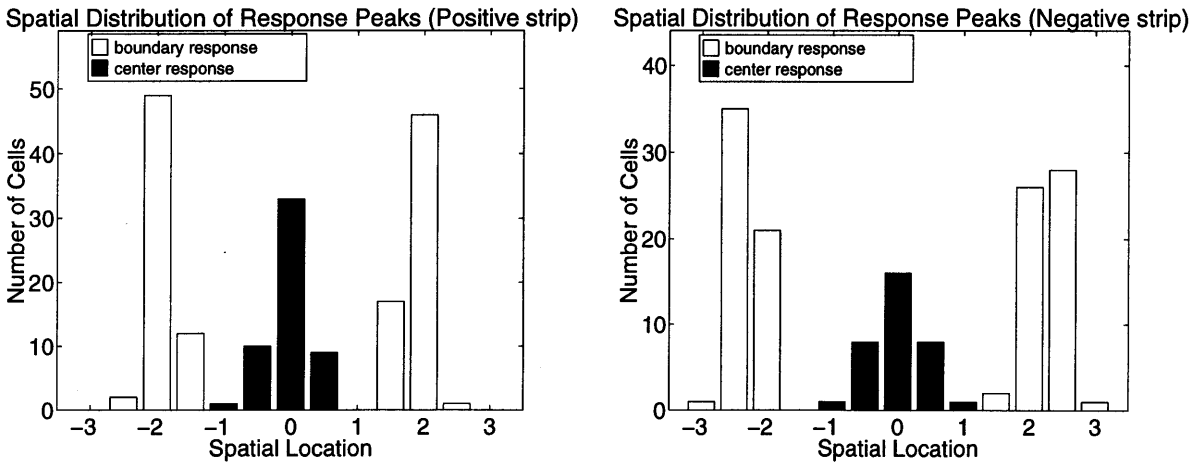

Fig. 12. The population histograms of the dispersion of the spatial locations of the interior response peaks and the boundary response peaks of all the cells (65 cells) tested with the strip stimuli. Results from different strip widths were combined. The spatial width of the strip was normalized to 4, so -2 and 2 were the locations of the strip boundary, and 0 should be the location of the center. The location of the interior response peak was the location of the maximum response peak from spatial location -1 to 1 within the figure. The location of the boundary response peak was the maximum response from -3 to -1 (boundary on the left), and 1 to 3 (boundary on the right), respectively. Cells that showed no statistically significant interior peak were not included in the interior peak cell count, but were included in the boundary response cell count in these histograms. As shown, some interior response peaks were deviated from the center, but as a whole population, the center was emphasized. Interestingly, the boundary response peaks were accurately localized, slightly biased inward in the positive strips, while they were strongly biased outward to -2.5 and 2.5 in the negative strip. This neural phenomenon might be related to the perceptual observation that the negative strip (with horizontal texture strips) appears visually to be wider than the positive strip (with vertical texture strips).

response peak. This potentially is a negative result for the medial axis hypothesis. Psychophysically, Kovacs and Julesz [36] also found that the sensitivity enhancement could be observed in a circle that was defined by random Gabor patches, but not with a white disk in a black background. Why, if the central response peak is signaling the location of the medial axis, would the central response not be observed in the black/white strip? One plausible explanation is that the medial axis signal might be sub-threshold and would reveal itself only when the receptive field of the cell was stimulated by the texture stimulus.

The results from the black ring experiment, although preliminary, were particularly illuminating. When the texture contrast borders were occluded by a black ring, the neurons' responses were markedly suppressed. This was not because the black-to-texture border signals were weaker than the texture-to-texture border signals. On the contrary, the texture disk in a black background with the same black-to-texture borders actually induced a much stronger response in the neurons than a texture disk surrounded by contrasting texture. One interpretation of this phenomenon is that when the ring was placed to occlude the texture contrast border, the ring became the figure, and the texture at the center was pushed back to become part of the texture background, and the neuron sitting at the center was no longer inside the figure, hence the suppression of the responses. This evidence, resonating with the result of Zipser et al.'s [26] frame and moat experiment and the ideas of amodal surface completion [40,41], shows that the interior enhancement requires the conjunction of a homogeneous surface and borders that belong to the surface. This condition is precisely the condition required for figure-ground computation as well as medial axis computation.

4. General discussion

4.1. Summary of findings

The main findings of the neurophysiological experiments presented in this paper are summarized below:

1. The initial responses (40–60 ms) of the neurons were characterized by filter responses to local features, while the later responses (80–200 ms) depended on contextual information and were possibly related to higher-order computations (Figs. 8 and 10).
2. V1 neurons' responses were enhanced when their receptive fields were inside a compact figure more so than when they were outside the figure, confirming, Lamme's [25] empirical findings (Fig. 10F).
3. The neurons were sensitive to the surface area of both texture figures and uniformly colored figures. In particular, the response decreased as the surface area or distance from borders increased (Figs. 17 and 18).
4. Orientation-preference of the cell was important in the later stage of V1 neurons' responses, particularly in the context of boundary detection. If the

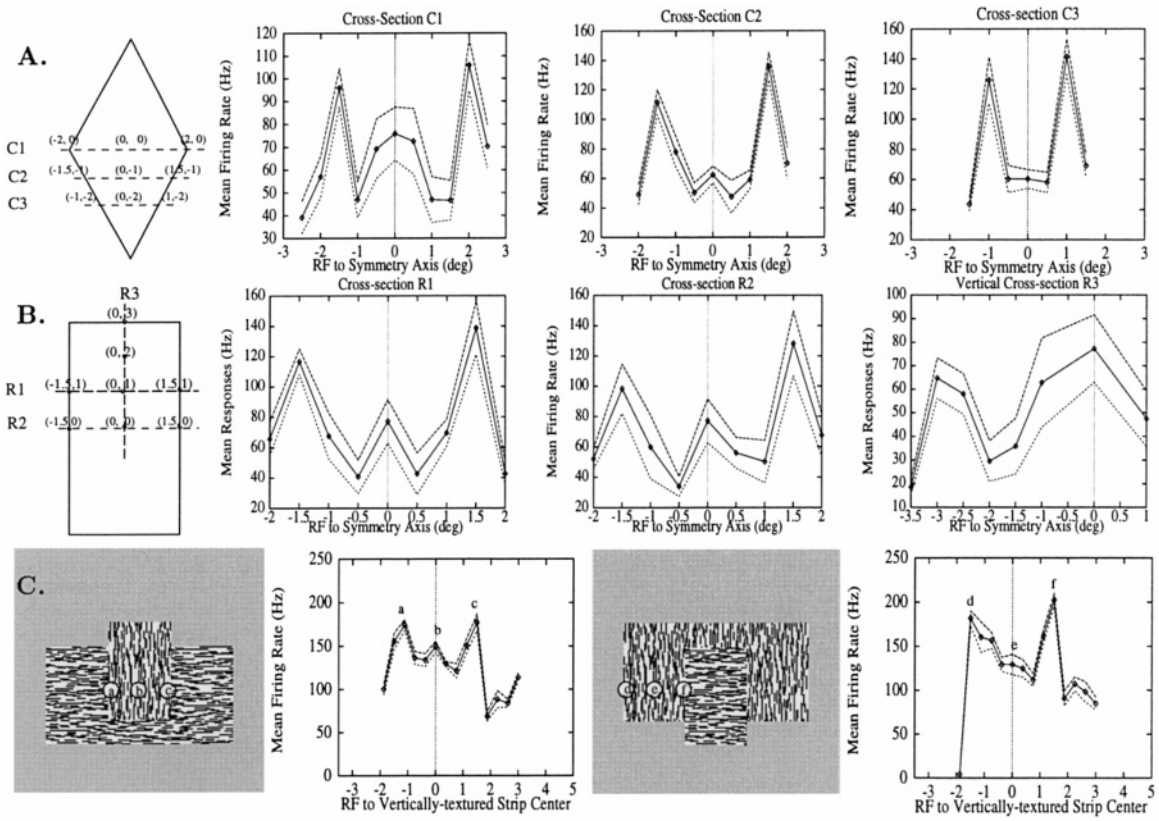

Fig. 13. The spatial response profiles of V1 neurons to more complex shapes, such as a diamond and a rectangle defined by texture contrast, were examined. The major axis of the shapes and the texture within the shapes were aligned with the preferred orientation of the cells. The shapes displayed here are for vertically oriented cells. The spatial response profiles of two neurons at the 100–200 ms post-stimulus time interval along different cross-sections of the shapes were shown. (A) Central peaks were observed in four out of 40 cells in response to an elongated diamond. In this particular cell, the central response peak was observed only at a certain location along the medial axis. Most of the cells showed distinct central peaks only for a particular range of figural size. A population of cells, each showing central peak of a particular width, can potentially encode the entire axis of a complex shape. (B) The response of a particular cell along cross-sections R1 and R2 shows statistically significant central response peaks along the mid-line of the rectangle, observed in five out of 15 cells. Other cells did not show any response peak within the figure. Along the longitudinal cross-section R3, there is also a response peak at the centroid of the shape. (C) Perceptually, the vertically-textured strip in the left figure is a vertical rectangle in front, while the same strip in the right figure is a part of the larger horizontal rectangle behind. This is the phenomenon of amodal surface completion described by Kaniza [40], and by Nakayama and Shimojo [41]. If the central peak is related to a medial axis representation computed with reference to the depth ordering relationship of objects in a visual scene, we would predict the central peak to show up in the left figure but not on the right figure. This is indeed the case for five out of 30 isolated cells we studied: central peaks were observed for the vertically textured strip of the left figure, but not for that of the right figure, three cells showed central peaks for both figures and two cells showed the reverse effect. An example of a cell's response is shown in the figure. The circles on the stimuli indicate the positions and the size of the receptive fields. The local stimuli at the corresponding sampling positions (e.g. a vs d, b vs e) were identical.

boundary was parallel to the preferred orientation of the cell, there was always a substantially stronger response at the boundary than other parts of the figure and background. When the texture features defining the boundary were not of the preferred orientation of the cell, the response at the boundary of the later stage was often stronger than the initial 'local feature detector' phase of the response (Fig. 10B).

5. In the non-parallel testing condition, the cell's response and the response enhancement were relatively uniform spatially (Fig. 10E and F).
6. In the parallel testing condition, apart from the boundary and interior enhancement, there was an additional response enhancement at the axis of the figure in a subset of neurons (Figs. 8 and 10A,C and D). These central response peaks were also observed in the response to stimuli (SSQ^+ and SSQ^-) in the slant testing condition (Fig. 10E).
7. The boundary response was invariant with respect to the surface area of the figure, but the central response peak and the interior response were dependent on the width and the surface area of the figure respectively. Neurons usually exhibited a central response peak only for a narrow range of widths. There was also a dispersion of the 'axis-response' peaks around the true center, but the emphasis was on the center (Figs. 11 and 12).

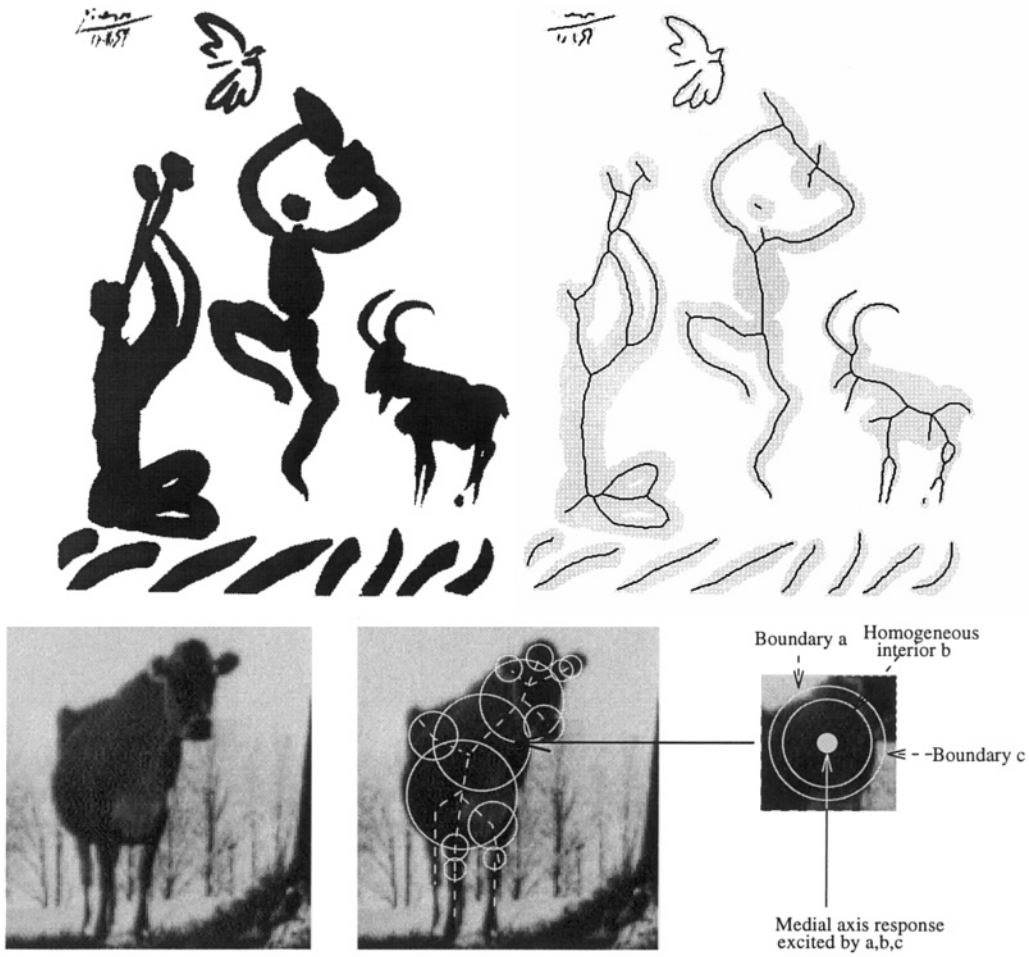

Fig. 14. (Top row): Piccaso's Rite of Spring (left) and the medial axis representation on the colored figures, generated by Ogniewicz's [67] algorithm (right). Figural coloring and medial axis computation are two critical visual routines important in shape recognition.(Bottom row): The figure illustrates how a cell may be constructed so that it fires when located on the medial axis of an object. The conjunction of three properties has to be present: at least two distinct segments of surface discontinuity on a disk of a certain radius and the homogeneity of surface qualities within an inscribing disk. Such a response is highly nonlinear but can be robustly computed.

However, central peak response was not observed in black/white strips (Fig. 19). While the boundary of the positive strip was accurately localized by V1 neurons, the locations of the boundary responses to the negative texture strip (i.e. with horizontal stripes) were more separated than the actual boundary locations (Fig. 12).

8. The interior response peak tended to emphasize the center of mass, rather than the entire medial axis of compact shapes such as diamonds and squares. Central response peaks were found along the entire axis only for elongated strips and rectangle (Fig. 13).
9. The response within a texture surface decreased relative to the response at the boundary of the surface over time, resulting in an increased localization and relative enhancement at the boundary. In contrast, the response within a uniformly colored figure increased relative to the boundary response over time, resulting in a spreading of border contrast signals toward the interior of the figure (Fig. 18).
10. The coexistence of a homogeneous surface and the borders belonging to the surface is necessary to produce the interior enhancement. Removing the borders that belong to the surface, e.g. occlusion by a black ring, eliminated the interior enhancement effect (Fig. 13(C), and Fig. 20).

4.2. Mechanistic interpretations

There are several approaches to interpreting these data. From a mechanistic point of view, it is known that extensive horizontal axonal collaterals in the superficial layers of V1 provide a medium for excitatory and inhibitory interaction among V1 neurons [42–44]. Various kinds of contextual phenomena, such as side-stopping [19] and pop-out [23] have been attributed to these intracortical connections. Various explicit mechanistic models [29,44] have been proposed using lateral

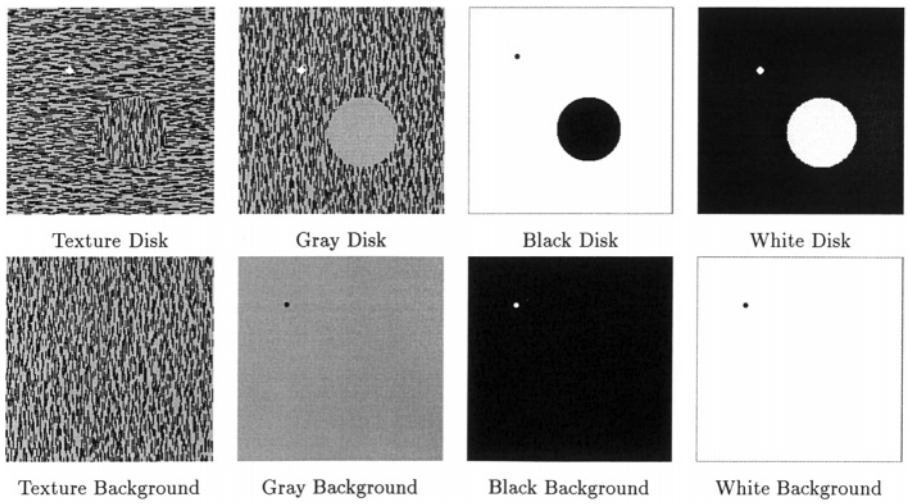

Fig. 15. In this experiment, 42 V1 neurons with receptive fields ranging from 0.8–1.5° were placed at the center of the texture disks or uniformly colored disks: white disk in a black background, black disk in a white background, gray disk in a white or black or texture background. Disks of three sizes (diameter 2, 4, 7°) were tested. The dot illustrated the position of the fixation spot.

inhibitory and excitatory interactions to account for these known effects, although feedback from extrastriate cortices has also been implicated in mediating at least the 'pop-out' phenomenon. In these models, the action of the surround is a function of the relative contrast between the center stimulus and the surround stimulus: when the contrast of the center stimulus is weaker than that of the surround, the surround influence tends to be facilitatory, whereas a center stimulus of strong contrast would tend to be suppressed by the surround units of similar orientation tuning. If properly constrained and parameterized, could these contextually dependent surround inhibition and facilitatory mechanisms account for the phenomena that we observed?

First, the decrease in the interior enhancement with the increase in the surface area of the texture figure could be understood from known lateral inhibition of the surrounding neurons with similar orientation tuning. The spatial extent ($1/e$ width of a Gaussian inhibitory kernel) of this lateral inhibition in awake monkeys was measured to be large as 3° at 3° eccentricity (Fig. 18) and 5–7° at 5° eccentricity (Fig. 17, see also ref. [26]). We found, through computer simulation, that a network with recurrent lateral inhibition of such a large spatial extent can produce the uniform and asymmetric enhancement within the figure that was observed under the non-parallel condition (Fig. 10F). However, lateral inhibition from units of similar orientation preference in the surround cannot account for the drastic interior enhancement observed in the non-orientation selective neurons by Lamme [25], because cells in the non-orientation selective 'channel' should respond equally well to the foreground and background texture. Moreover, it cannot account for the size-dependent interior enhancement observed in uniformly colored figures because these figures are defined only by luminance or color surface and contours with no oriented local texture features in either the background or the foreground to excite the orientation-selective units.

Therefore, a lateral excitatory mechanism is required. Stemmler et al. [29] and Somers' models do provide a mechanism for facilitation of the surround when the center stimulus is low in contrast relative to the surround texture. This mechanism may be generalized so that a strong border response in the surround likewise can facilitate or excite a weak center, effecting the enhancement in uniformly colored figures as well as negative texture figures. Our observation that the texture border signals tend to sharpen and the uniformly colored border signals tend to spread (Fig. 18) further suggests that both lateral excitation and lateral inhibition are involved. Although lateral excitation is bi-directional, it is possible to obtain relative interior enhancement simply because the signals are converging from the borders inside the figure and are diverging from the borders outside the figure.

However, we found the following four pieces of evidence particularly difficult to explain with lateral excitatory and inhibitory mechanisms alone. First is the result of the black ring experiment, similar in conclusion to Zipser et al.'s [26] stereo frame and moat experiment, which seems to suggest that the coexistence of a homogeneous surface with the occluding borders belonging to the surface is required to cause interior enhancement within the surface. In our experiment, the black ring produced a very strong border effect but did not induce interior enhancement, even though the same black-to-texture border signal shown as a texture disk in a black background (Fig. 20) could induce a strong

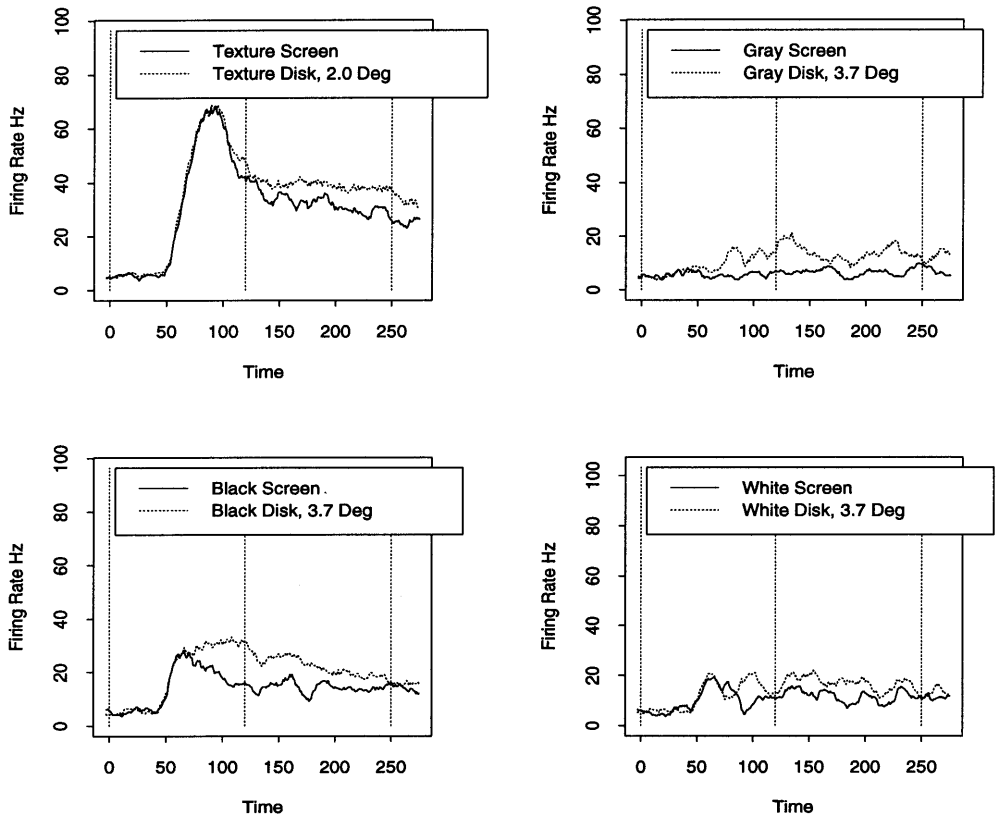

Fig. 16. The average response of 35 neurons in parafoveal area V1 (eccentricity 5–6°) at the center of texture disks or uniformly colored disks showed a significant enhancement over the response to the background stimuli. The 'background' stimulus was a full screen stimulus composed of the same cue that made up the corresponding disks. For example, the 'background' stimulus to compare with the white disk stimulus was the white screen. Between trials, the monitor screen was gray. Therefore, both the onset of a white disk or white screen presented a temporal luminance edge to the cells' receptive field. Six other neurons sampled at eccentricity 3° showed a similar but smaller effect. These profiles were smoothed by a 15 ms running average.

interior enhancement effect. These observations, together with Lee et al.'s [45] shape from shading experiment which showed that a convex 'pop-out' object can induce interior enhancement but a concave 'pop-in' hole cannot, suggests that the interior enhancement is contingent on the figure-ground percept and is not due simply to propagation of the border signals alone. Secondly, although the axis enhancement in the positive texture figures might be explained in terms of local lateral inhibition and dis-inhibition or integration beyond the zone of an inhibitory region (see ref. [20]). it is difficult to see how a response peak could emerge precisely at the center of a region where there was absolutely no response at all for the first 80 ms (Fig. 6). The clustering of the response peaks around the center of the figure (Fig. 12) also is unlikely a random effect of 'blind' lateral inhibition and excitation mechanisms. Thirdly, the evidence that the later response at the boundary could be stronger than the initial response of the neurons when the boundaries, but not the local features, were of the preferred orientation of the cells, could not be understood in terms of lateral inhibition either. Lateral inhibition can produce 'relative enhancement' in response but is not known to produce a greater absolute response in the later stage than the initial 'filter' response, which tends to be the strongest in ordinary situations. This suggests a more 'symbolic' description of the boundary might have emerged, possibly after articulation and elaboration within the local circuit under the influence of extrastriate feedback. Finally, the fact that the substantial part of the interior enhancement signals emerged around 80–100 ms after stimulus onset allows for feedback from extrastriate cortices, including IT [46], to impose more abstract and global constraints to the processing of information in V1.

Consistent with this feedback hypothesis, psychophysical studies [47] and PET and NMR imaging studies [48] have also shown that V1 activity can be modulated by various top-down influences, including mental imagery. Recent neurophysiological studies have shown that V1 neurons can be modulated by attention [49,50] and preliminary deactivation experimental results also suggested that some context-sensitive surround effects can be eliminated by deactivating V2 [51] or lesioning extrastriate cortical areas beyond V2 [52]. Given that V2 and even IT neurons start responding to the stimulus

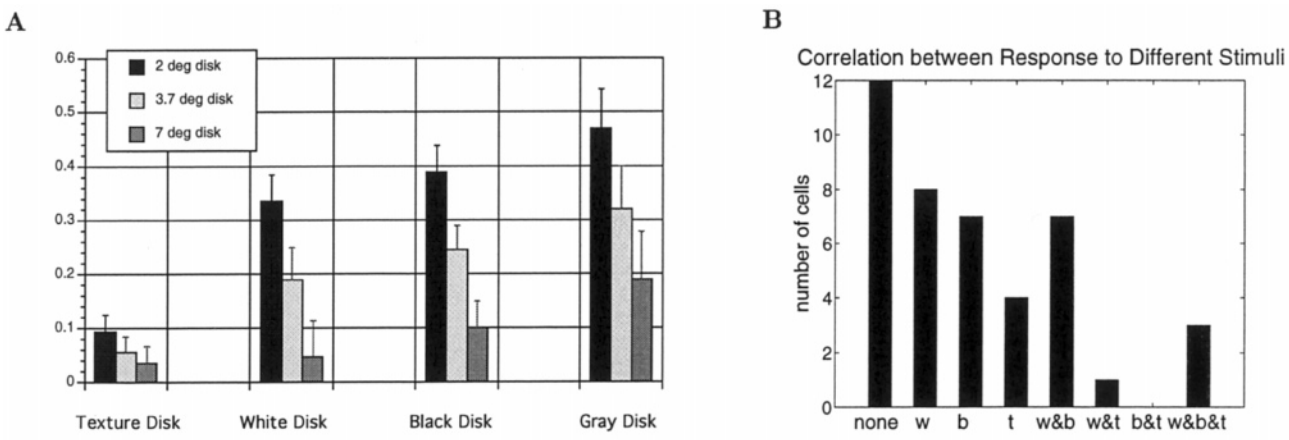

Fig. 17. (A.) The figure-ground enhancement modulation was the ratio $(A-B)/(A+B)$ where $A=$ the response to the center of a disk (row 1, Fig. 16) and $B=$ the response to a homogeneous screen (background) (row 2, Fig. 16), e.g. $A=$ the response to a white disk and $B=$ the response to a white screen. The mean figure-ground modulation indices showed a similar increase in modulation with the decrease in figural size. Note that figural enhancement was observed even for the 7° disk which was much bigger than the receptive fields. The confidence bar indicates a 95% confidence interval for the population mean (35 neurons at eccentricity 5–6°) within the 120 to 250 ms time window. (B.) The population distribution ($N=42$) of co-occurrence of interior enhancement for black disk (B), white disk (W) and texture disk (T). A cell is considered to exhibit positive interior enhancement to disks of a certain cue if it showed a statistically significant decrease in response when the diameter of the disk increased from 3.7 to 7° visual angle. The graph shows three out of 42 neurons were sensitive to disks of all three cues, eight were sensitive to at least two cues simultaneously, and 12 cells did not show enhancement for any of the three cues.

60–80 ms after stimulus onset [53], the long latency of the various phenomena we observed in V1 further suggests feedback likely plays a major role in mediating these effects.

4.3. Computational interpretations

Considering the data from a computational perspective might allow us to gain a deeper insight as to what functional purposes are served by these intracortical and feedback mechanisms. The computational perspective provides an alternative approach, but is not necessarily orthogonal to the traditional mechanistic interpretation.

One plausible computational interpretation to the enhanced neural activities within the figure is that V1 neurons are carrying out various kinds of visual routines. One of these routines is called 'coloring' or the labelling of all the 'pixels' that belong to the same surface of the object, separated from the background. Coloring a region takes time: Paradiso and Nakayama [54] showed that the percept of a white disk forms in stages over 50–100 milliseconds, propagating in from the edges, and that it can be interrupted by masks of different shapes. Our data show that the activities of interior enhancement arise around 60 ms for uniformly colored figures, and around 80 ms for texture figures. This is consistent with the psychophysical data. Coloring is more rapid relative to filling-in [24] which typically took seconds. There are now at least two possible related mechanisms to 'color' a region. One is to link activity of already active cells, as Gray and Singer's [55] data might suggest, using synchrony of spikes to represent the presence of a single large percept. The other, as data from this study might suggest, is to dedicate a cell or part of a cell's activity to signaling that specific elementary shapes, such as the disks, are part of a single surface not cut up by boundaries. This activity might be a part of the enhanced responses of the neurons within the figure. Furthermore, the locus of the centers of such disks could signal the location of the medial axis.

One purpose of coloring is to enhance the saliency of the figure for recognition. Figural saliency is controlled by many factors. The pop-out effect [35,56] is a result of automatic bottom-up analysis of the data. Saliency can also be enhanced by feedback as in the interactive activation models [2] in which extrastriate cortical neurons, presumably representing higher order internal representation, are bi-directionally connected to early cortical neurons that are sensitive to low and intermediate level local features. The internal object representations have been shown psychophysically to be able to group local features together to enhance local feature detection [57]. The activation of local features (e.g. the old man's ear in Fig. 1) would lead to the activation of the whole percept (e.g. the old man), which in turn would feedback to V1 to make the figure more salient to focus further computation. The enhanced activities then reflect partly the coloring process and partly the co-activation of the visual neurons at different levels of processings. The 80 ms latency in the enhanced activities may reflect the continuous articulation and elaboration of the visual analysis in V1 under the feedback from more abstract internal representations.

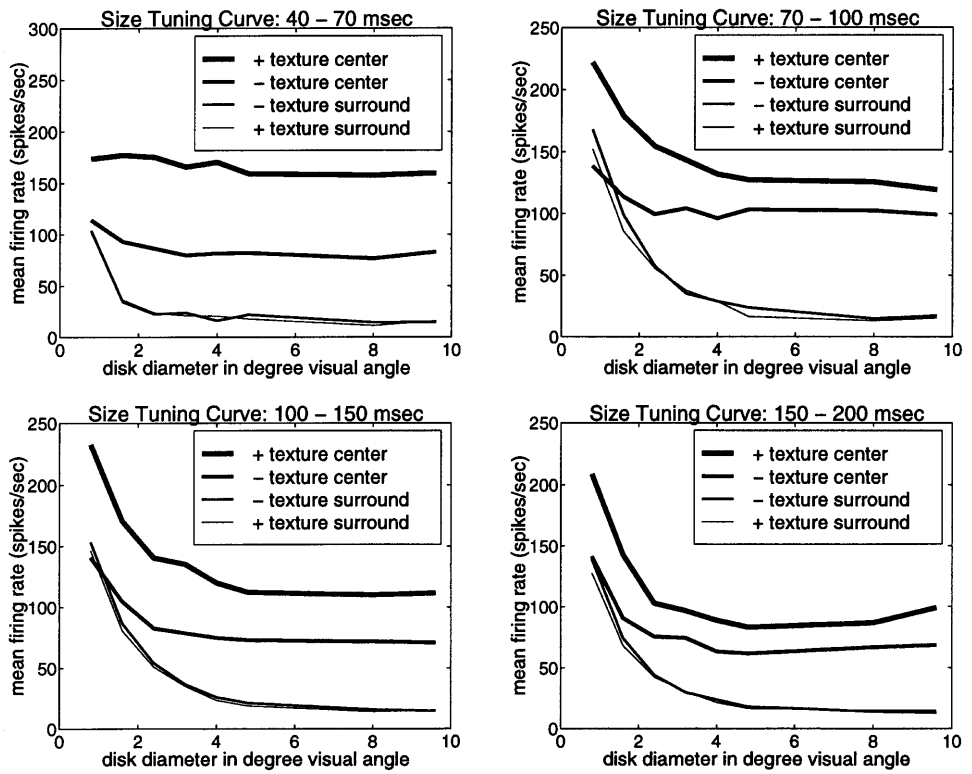

Fig. 18. The average firing rates of 20 neurons, responding to texture disks surrounded by contrasting texture and gray disks surrounded by texture of different diameters at different time windows after stimulus onset. The textures within the texture disks (texture center) or outside the gray disks (texture surround only) were parallel to the preferred orientation of the cell (+) or orthogonal to the preferred orientation of the cell (−). The response to both types of disks was sensitive to the distance away from the boundary. The response within the texture disks (texture center) was sensitive to the orientation of the texture inside the figure, but the responses within the gray disks (texture surround) were insensitive to the texture outside the figure, as shown in the population average. The activities within the texture disks contracted toward the boundary, while the activities within the gray disks showed signals spreading away from the boundary, revealing two possible concurrent processes of boundary detection and surface interpolation.

Only when a figure has become 'colored', i.e. separated from the background, can one begin to compute properties of an object's shape. These properties such as aspect ratios, parts and symmetries may in turn suggest the identity of the visible object. The medial axis or the skeleton transform are fundamental to structural and grammatical analysis of the shape of visual objects. It is critical for the construction of the hierarchical and modular 3D model of an object as proposed by Marr , or the representation of an object using structural primitives and spatial relationships [12,14]. Even though current neurophysiological and psychophysical evidence favors a more view-based approach to object recognition [58] in which objects are represented by combination of multiple views or aspects, there might still be a need for constructing the 2.5D sketch, and making use of a hierarchical, structural and grammatical approach to represent objects because such representations tend to be more efficient. The psychological evidence [36,38] and the neurophysiological evidence presented in this paper lends further support to the plausibility for such a structural analysis.

Another dimension of the computational interpretation in addition to image segmentation and figure-ground segregation is the construction of what Marr called the 2.5D surface sketch, a representation of the visible surface in depth. The process is also called surface reconstruction in computer vision. Computer scientists have proposed for several years now that boundary detection cannot be successful with sophisticated local edge detectors alone, but is best accomplished interactively with surface reconstruction [59–61] and this computational architecture has been explicitly proposed in preliminary forms for the visual cortex in luminance boundary detection [62] and in texture segmentation [60]. Grossberg [63] has also arrived at a similar neural network structure from a more psychological perspective. These computational models involve two concurrent and interactive processes. One is the boundary sharpening process that involves refinement of boundary signals as the surface representation is being formed, and the other is a surface interpolation process that involves the spreading of surface signals from the boundary. In our data, the contraction of signals within a texture figure toward the borders and

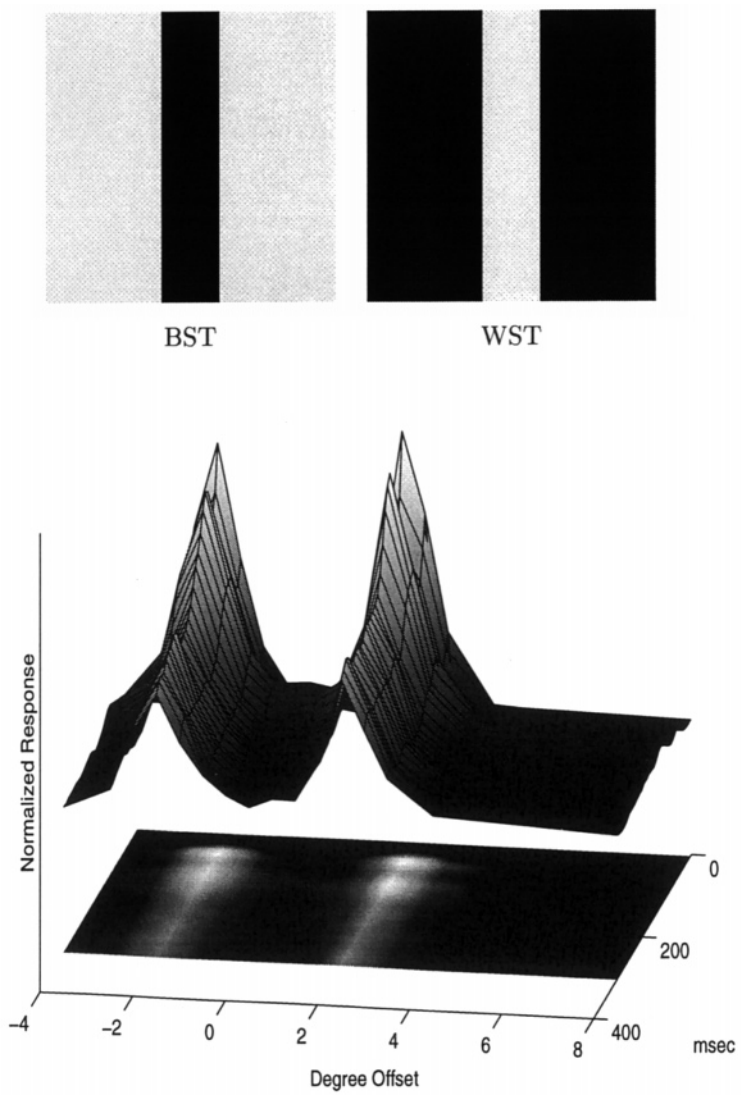

Fig. 19. Average spatial response profile of a population of 12 complex cells responding to BST (black strip) and WST (white strip), showing almost instantaneous boundary formation (40 ms), and a slight spreading of the border excitation signals inside and outside the strip during the resurgent period (80–120 ms after stimulus onset). The response is the combined response of BST and WST using the same averaging procedure described in Figs. 8 and 10. Central response peaks could not be observed in these black and white strip stimuli.

the spreading of border signals within the uniformly colored figure (Fig. 18) possibly reflect these concurrent processes of boundary detection and surface interpolation. Since V1 has heavy intracortical connections in the superficial layers, and the neurons are precisely localized in space and tuned in orientation, color, disparity and other cue modalities, it is an ideal 'working area' to solve these 1D and 2D geometry problems and to represent explicitly the boundary contours and the 2.5D sketch in a retinotopic map. Given that segmentation and surface reconstruction cannot be fully accomplished without figure-ground distinction and object recognition, all levels of visual processing would be involved in constructing these representations.

4.4. Conclusion

Our main thesis is that the later part of V1's neural activities reflect their participation in multiple visual routines and higher level visual processing's. The temporal progression of activities of V1 neurons reflects the gradual involvement of V1 in successively higher levels of computations (Fig. 21). Within this framework, the spike train of a neuron reflects the 'aggregate' of all activities a neuron is engaged in. Imagine each visual routine engages a neuron in a particular intracortical and intercortical neural circuit, as a neuron participates in more visual routines and gets involved in more neural circuits, it will become more active. Our data

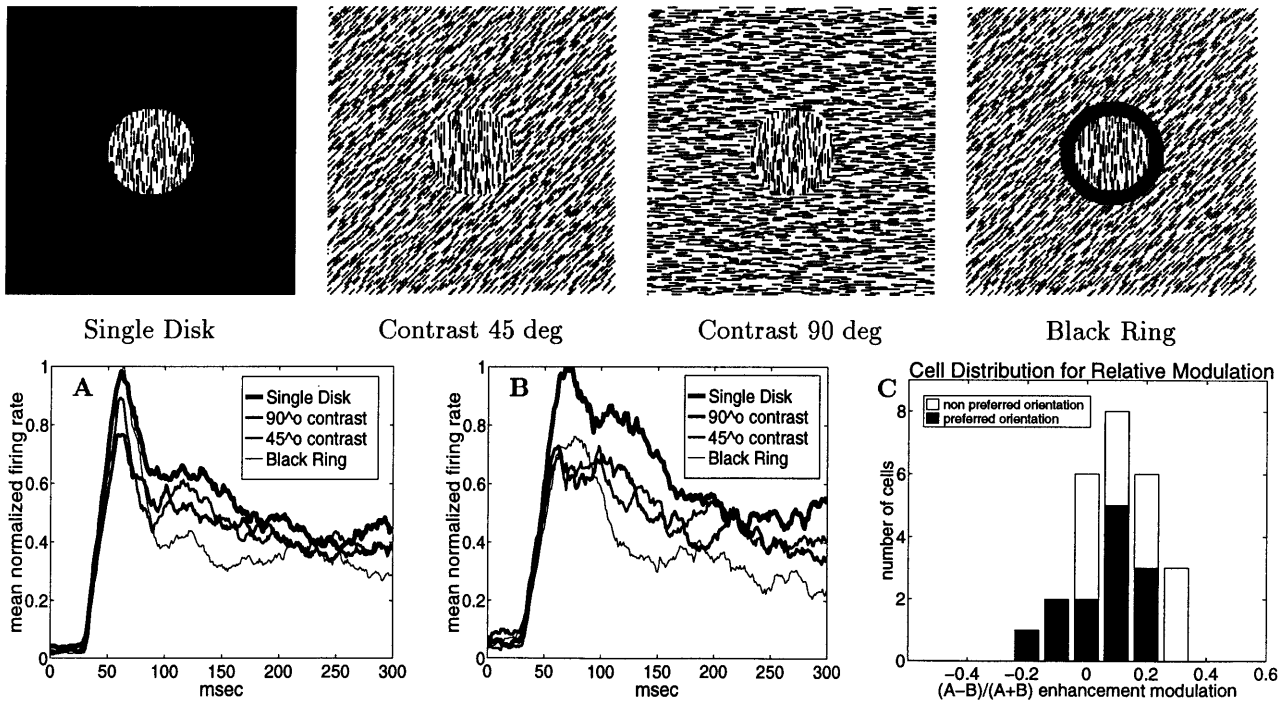

Fig. 20. Nine V1 neurons of one monkey were tested with the stimuli shown. The receptive fields of the neurons were about 1° in diameter at 4° visual eccentricity. The receptive fields were placed at the center of the texture disk (2° visual angle in diameter). (A., B.) The average normalized temporal responses of the neurons to the four stimuli were shown. In A, the texture within the disk was parallel to the preferred orientation of the cells. In B, the texture within the disk was orthogonal to the preferred orientation of the cells. In both A and B, the single disk evoked the greatest response, producing a greater initial burst of response particularly in case B. Responses to the texture contrast disks were relatively independent of the degree of texture contrast, i.e. invariant to the orientation of the texture outside the disk. Adding a black ring to occlude the texture contrast borders suppressed the response to the texture contrast disks at the later stage of the response. The histogram (C.) shows the distribution of the neurons using the interior enhancement ratio $(A-B)/(A+B)$, where A was the neural response to the 45° texture contrast disk and B was the neural response to the black ring stimulus. When the disk texture was of the preferred orientation, there was a significant bias towards positive enhancement, i.e. suppression by the black ring. When the disk texture was of the non-preferred orientation, the black ring suppression was even more significant.

seems to suggest individual V1 neurons can indeed participate in multiple computations and encode multiple representations. The results of these computations and representations likely become multiplexed in the spike train, possibly using synchronized spikes in the population of neurons or complex neural assemblies [64].

Our proposal, that V1 is engaged in many levels of visual analysis through intracortical and feedback connections, is a significant departure from the classical feed-forward views on the nature of information processing and the functional role of V1 [11]. Classical ideas going back to Hubel and Wiesel attempt to interpret all neuronal responses as feature detectors, modulated by various contextual factors. In the case of V1, this amounts to filters with various extra receptive field enhancements and suppressions. This framework is so broad that almost all effects can be coerced into it, using complex higher order effects such as dis-inhibition and integration beyond lateral inhibition. We find that although it can be stretched to account for most of the data, we don't believe it is a very simple or parsimonious one, nor does it account for some of the more striking experimental results. But the computational approach, stemming from Marr, takes a radically different view. It attempts to identify the visual structures that must be computed and then see if single cell responses indicate that they are being computed, taking an agnostic view about whether cells are individually signalling features or acting in complex assemblies. The latter approach has been successful in relating V2 responses to Gestalt theory [65,66]. It seems to us that the computational interpretation is the simpler and preferable alternative.

Taken together, the data presented in this paper and others [25–28,45] suggest that the V1 is not just a module for computing local features, but possibly serves as a high resolution buffer or visual computer to perform all computations that integrate global information with spatial precision. The intricate intracortical circuitry in V1, together with the recurrent extrastriate cortical feedback, allows V1 to participate in many levels of visual analysis and to represent many kinds of higher level structural information which are critical to recognition.

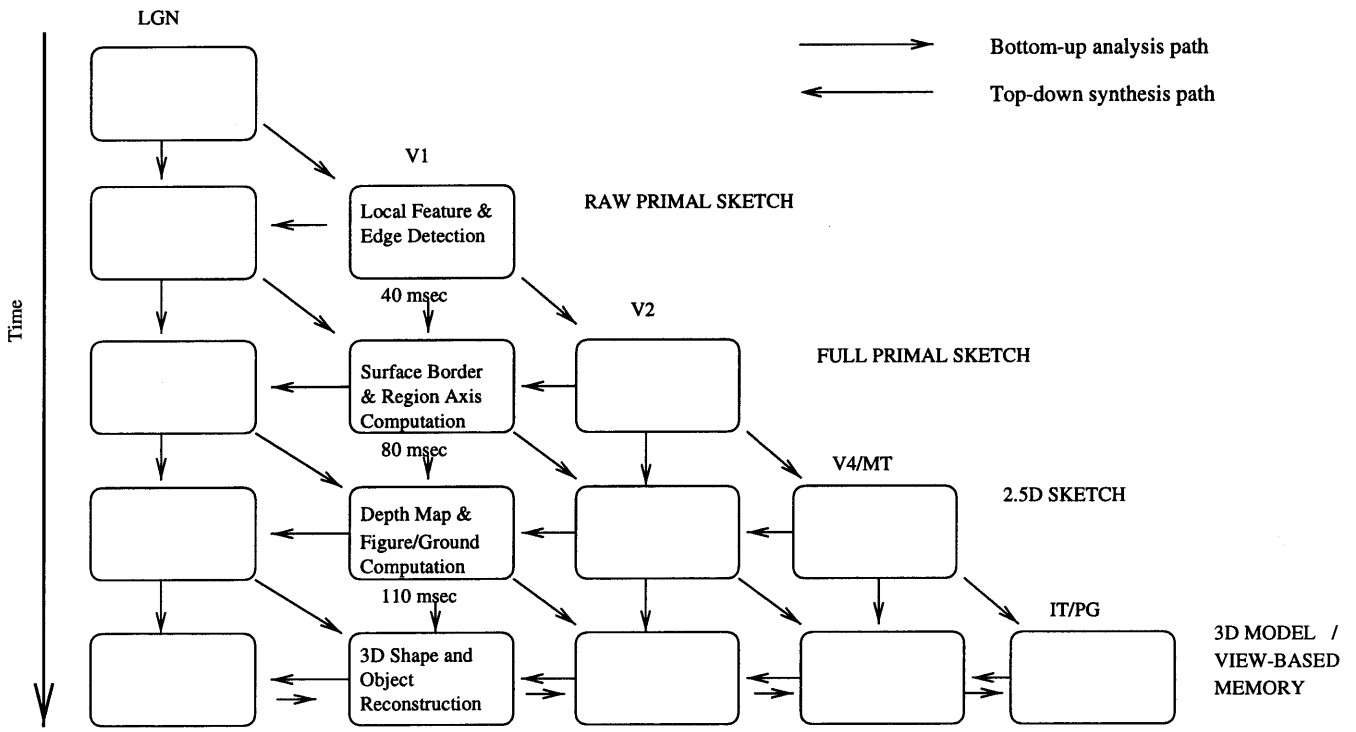

Fig. 21. This schematic diagram illustrates our ideas on how V1 becomes engaged in different levels of visual processing at different times post-stimulus onset. From 40–60 ms, the responses of the neurons are characterized as local feature and edge detectors. From 80 ms onwards, cue-invariant boundary signals are computed and represented. From 90–110 ms, interior enhancement (figure-ground signals) and axis signals are produced. Another study by Lee et al. [45] shows that V1 neurons are sensitive to 3D shape from shading information 110 ms after stimulus onset. Image segmentation, figure-ground, shape computation and object recognition in this framework occur concurrently and interactively in a constant feed-forward and feedback loop that involves the entire hierarchical circuit in the visual system. Signals of higher level visual representations, such as a 2.5D surface sketch, 3D model or view-based object memory, are likely reflected in the later part of V1's activities.

Acknowledgements

The experiments reported here were conducted at Peter Schiller's Lab at MIT. We are grateful to P.H. Schiller for encouragement and advice, to E. Cassidente, W. Slocum and K. Zipser for technical assistance, and to D. Pollen, I. Kovacs, C.E. Ho and many colleagues at Harvard, MIT and CMU including C. Olson, M. Behrmann, S.C. Zhu, S. Gettner, J. Goodridge, K. Rearick, A. Tolias, T. Moore and J. Mazer, for helpful discussion and reading many versions of the manuscript. This research is supported by a McDonnell-Pew Grant to T.S. Lee, a NSF grant DMS-93-21266 to D. Mumford, a NSF training fellowship to R. Romero, a grant from The Netherlands Organization for Scientific Research to V.A.F. Lamme and a NIH grant EY00676 to P.H. Schiller.

References

[1] Ullman S. Visual routines. Cognition 1984;18:97–159.
[2] McClelland JL, Rumelhart DE. An interactive activation model of context effects in letter perception. Part I: an account of basic findings. Psychol Rev 1981;88:375–407.
[3] Grossberg S. Competitive learning: from interactive activation to adaptive resonance. Cogn Sci 1987;11:23–63.
[4] Mumford D. On the computational architecture of the neocortex II. Biol Cybern 1992;66:241–51.
[5] Mumford D. Neuronal architectures for pattern-theoretic problems. In: Koch C, Davis JL, editors. Large-Scale Neuronal Theories of the Brain. Cambridge, MA: MIT Press, 1994:125–52.
[6] Ullman S. Sequence seeking and counters/reams: a model for bi-directional information flow in the cortex. In: Koch C, Davis J, editors. Large-Scale Theories of the Cortex. Cambridge, MA: MIT Press, 1994:257–70.
[7] Dayan P, Hinton GE, Neal RM, Zemel RS. The Helmholtz machine. Neural Comput 1995;7(5):889–904.
[8] Daugman JG. Uncertainty relation for resolution in space, spatial frequency, and orientation optimized by two-dimensional visual cortical filters. J Opt Soc Am 1985;2(7):1160–9.
[9] Pollen DA, Gaska JP, Jacobson LD. Physiological constraints on models of visual cortical functions. In: Rodney M, Cotterill J, editors. Models of Brain Function. New York: Cambridge University Press, 1989:115–35.
[10] Blum H. Biological shape and visual science J. Theor Biol 1973;38:205–87.
[11] Marr D. Vision. New York: WH Freeman, 1982.
[12] Biederman I. Recognition-by-components: a theory of human image understanding. Psychol Rev 1987;94(2):115–47.
[13] Pizer SM, Oliver WR, Bloomberg SH. Hierarchical shape description via the multi resolution symmetric axis transform. IEEE Trans PAMI-9, 1987:4.

[14] Zhu SC, Yuille AL. Forms: a flexible object recognition and modelling system. Int J of Comp Vis 1995;187–212.
[15] Olshausen BA, Anderson CH, Van Essen DC. A multiscale dynamic routing circuit for forming size- and position-invariant object representations. J Comput Neurosci 1995;21(1):45–62.
[16] Maffei L, Fiorentini A. The unresponsive regions of visual cortical receptive fields. Vis Res 1976;16:1131–9.
[17] Fries W, Albus K, Creutzfeldt OD. Effects of interacting visual patterns on single cell responses in cat's striate cortex. Vis Res 1977;17:1001–8.
[18] Nelson JI, Frost B. Orientation selective inhibition from beyond the classical visual receptive field. Brain Res 1978;139:359–65.
[19] Born RT, Tootell RBH. Single-unit and 2-deoxyglucose studies of side inhibition in macaque striate cortex. Proc Natl Acad Sci 1991;88:7071–5.
[20] Li CY, Li W. Extensive integration field beyond the classical receptive field of cat's striate cortical neurons-classification and tuning properties. Vis Res 1994;34:2337–55.
[21] Sillito AM, Grieve KL, Jones HE, Cudeiro J, Davis J. Visual cortical mechanisms detecting focal orientation discontinuities. Nature 1995;378(6556):492–6.
[22] Gilbert CD, Wiesel TN. The influence of contextual stimuli on the orientation selectivity of cells in the primary visual cortex of the cat. Vis Res 1990;30:1689–701.
[23] Knierim JJ, Van Essen DC. Neuronal responses to static texture patterns in area V1 of the alert macaque monkey. J Neurophysiol 1992;67:961–80.
[24] De Weerd P, Gattas R, Desimone R, Ungerleider LG. Center-surround interactions in area V2/V3; A possible mechanism for filling in? Soc Neurosci Abstr 1993;19:27.
[25] Lamme VAF. The neurophysiology of figure-ground segregation in primary visual cortex. J Neurosci 1995;10:649–69.
[26] Zipser K, Lamme VAF, Schiller PH. Contextual modulation in primary visual cortex. J Neurosci 1996;16(22):7376–89.
[27] Gilbert CD, Das A, Ito M, Kapadia M, Westheimer G. Spatial integration and cortical dynamics. Proc Natl Acad Sci USA 1996;9S:615–22.
[28] Levitt JB, Lund JS. Contrast dependence of contextual effects in primate visual cortex. Nature 1997;387:73–6.
[29] Stemmler M, Usher M, Niebur E. Lateral interactions in primary visual cortex: a model bridging physiology and psychophysics. Science 1995;269:1877–80.
[30] Somers DC, Todorov EV, Siapas AG, Toth LJ, Kim DS, Sur M. A local circuit integration approach to understanding visual cortical receptive fields. Cerebral Cortex 1997 (in press).
[31] Robinson DA. A method of measuring eye movement using a scleral search coil in a magnetic field. IEEE Trans Biomed Electron IOI, 1963;131.
[32] Haenny PR, Schiller PH. State dependent activity in monkey visual cortex I single cell activity in V1 and V4 on visual tasks. Exp Brain Res 1988;69:225–44.
[33] Gallant JL, Van Essen DC, Nothdurft HC. Two-dimensional and three-dimensional texture processing in visual cortex of the macaque monkey. In: Papathomas TV, Chubb C, Gorea A, Kowler E, editors. Early Vision and Beyond. Cambridge, MA: MIT Press, 1994:89–98.
[34] DeValois RL, Albrecht DG, Thorell LG. Spatial frequency selectivity of cells in macaque visual cortex. Vis Res 1982;22:545–59.
[35] Julesz B. Experiments in the visual perception of texture. Sci Am 1975;232:34–43.
[36] Kovacs I, Julesz B. Perceptual sensitivity maps within globally defined visual shapes. Nature 1994;870:644.
[37] Crowley J, Parker AC. A representation for shape based on peaks and ridges in the difference of low-pass transform. IEEE Trans Pattern Recog Mach Intell 1984;6(2):156–70.
[38] Burbeck CA, Pizer SM. Object representation by cores: identifying and representing primitive spatial regions. Vis Res 1995;35(13):1917–30.
[39] Livingston M, Hubel DH. Specificity of intrinsic connections in primate primary visual cortex. J Neurosci 1984;4:2830–5.
[40] Kaniza G. The Organization of Vision. New York: Praeger, 1979.
[41] Nakayama K, Shimojo S. Experiencing and perceiving visual surfaces. Science 1992;257:1357–62.
[42] Livingston M, Hubel DH. Anatomy and physiology of a color system in the primate visual cortex. J Neurosci 1984;4:309–56.
[43] Ts'o D, Gilbert CD, Wiesel TN. Relationships between horizontal interaction and functional architecture in cat striate cortex as revealed by cross-correlation analysis. J Neurosci 1986;6(4):1160–70.
[44] Lund JS, Wu Q, Hadingham PT, Levitt JB. Cells and circuits contributing to functional properties in area V1 of macaque monkey cerebral cortex: bases for neuroanatomically realistic models. J Anat 1995;87:563–81.
[45] Lee TS, Mumford D, Romero R, Tolias A, Moore T. Sensitivity of V1 neurons to shape from shading. Invest Opt Vis Sci 1997;38:459.
[46] Rockland KS, Van Hoesen GW. Direct temporal-occipital feedback connections to striate cortex (V1) in the macaque monkey. Cereb Cortex 1994;4(3):300–13.
[47] Ishai A, Sagi D. Common mechanisms of visual imagery and perception. Science 1995;268:1772–4.
[48] Kosslyn S, Thompson WL, Kim IJ, Alpert NM. Topographical representations of mental images in primary visual cortex. Nature 1995;378:496–8.
[49] Motter BC. Focal attention produces spatially selective processing in visual cortical areas V1, V2, V4 in the presence of competing stimuli. J Neurophysiol 1993;70(3):909–19.
[50] Press WA, Knierim JJ, Van Essen DC. Neuronal correlates of attention to texture patterns in macaque striate cortex. Soc Neurosci Abstr 1994;20:838.
[51] James AC, Hupe JM, Lomber SL, Payne B, Girard P, Bullier J. Feedback connections contribute to center surround interactions in neurons of monkey area V1 and V2. Soc Neurosci Abstr 1995;21:359.10.
[52] Lamme VAF, Zipser K, Spekreijse H. Figure-ground signals in V1 depend on extrastriate feedback. Invest Opt Vis Sci 1997;38:969.
[53] Bullier J, Nowak LG. Parallel versus serial processings: new vistas on the distributed organization of the visual system. Curr Opin Biol 1995;5:497–503.
[54] Paradiso MA, Nakayama K. Brightness perception and filling-in. Vis Res 1991;31:1221–36.
[55] Gray CM, Singer W. Stimulus-specific neuronal oscillations in orientation columns of cat visual cortex. Proc Natl Acad Sci USA 1989;86:1698–702.
[56] Treisman A. Perceptual grouping and attention in visual search for features and for objects. J Exp Psychol Hum Percept Perform 1982;8:194–214.
[57] Behrmann M, Zemel R, Mozer M. Object-based attention and occlusion: evidence from normal subjects and a computational model. J Exp Psychol Human Percept Perform 1997 (in press).
[58] Logothetis N, Paul J, Bulthoff HH, Poggio T. Viewer-dependent object recognition by monkeys. Curr Biol 1994;4(5):401–14.
[59] Blake A, Zisserman A. Visual Reconstruction. Cambridge, MA: MIT Press, 1987.
[60] Lee TS. A Bayesian framework for understanding texture segmentation in the primary visual cortex. Vis Res 1995;35(18):2643–57.
[61] Belhumeur PH. A Bayesian approach to binocular stereopsis. Int J Comput Vis 1996;1–26.

[62] Koch C, Marroquin J, Yuille AL. Analog 'neuronal' networks in early vision. Proc Natl Acad Sci USA 1986;83:4263–7.
[63] Grossberg S. 3-D vision and figure-ground separation by visual cortex. Percept Psychophys 1994;55(1):48–120.
[64] Abeles M. Corticonics: Neural Circuits of the Cerebral Cortex. Cambridge, UK: Cambridge University Press, 1991.
[65] Baumann R, Zwan RVD, Peterhans E. Figure-ground segregation at contours: a neural mechanism in the visual cortex of the alert monkey. Eur J Neurosci 1997;9:1290–303.
[66] von de Heydt R, Peterhans E. Illusory contours and cortical neuron responses. Science 1984;224(4654):1260–2.
[67] Ogniewicz R. Skeleton-space: a multiscale shape description combining region and boundary information. Proc Conf Comput Vis Pattern Recog 1994;746–751.