A ROLE FOR STRIATE CORTEX IN SURFACE PERCEPTION

by

Karl Zipser

B.A., University of Chicago, Chicago (1991)

Submitted to the Department of Brain and Cognitive Sciences in Partial
Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 1995

Signature of Author______________________________
Department of Brain and Cognitive Sciences
September 6, 1995

Certified by______________________________
Peter H. Schiller
Professor, Department of Brain and Cognitive Sciences
Thesis Supervisor

Accepted by______________________________
Gerald E. Schneider
Chairman, Department Graduate Committee

A ROLE FOR STRIATE CORTEX IN SURFACE PERCEPTION

by

Karl Zipser

Submitted to the Department of Brain and Cognitive Sciences in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy in Neuroscience

Abstract

In this thesis I describe a series of experiments conducted on striate cortex (V1) of the awake, behaving monkey. The goal of these experiments was to examine the manner in which stimuli presented outside the receptive field (RF) of V1 neurons may influence the activity of these cells evoked by direct RF stimulation. I term this phenomenon *extra-RF modulation.* In the first chapter I describe experiments designed to study basic characteristics of extra-RF modulation, such as how extra-RF modulation may be evoked by different visual cues, and what is the spatial extent over which extra-RF modulation may operate in the visual field. Data collected in these experiments show that extra-RF modulation may indeed be evoked by diverse cues (binocular disparity, luminance, and color) in addition to orientation cues which have previously been studied. Furthermore, the data show that extra-RF modulation operates over a large spatial range (a diameter of 10° of visual angle centered on the RF, on average.) In the second chapter I describe experiments designed to study the functional role of extra-RF modulation in visual analysis. In these experiments extra-RF modulation was examined with displays of textured surfaces that may be seen as occlude by other structures. The results from these experiments raise the possibility that extra-RF modulation plays a role in the perception of visual surfaces. In the third chapter I describe experiments that probe the dynamic properties of extra-RF modulation.

Thesis Supervisor: Peter H. Schiller

Acknowledgments

I thank Peter Schiller for his unwavering support, encouragement, and guidance. I thank Victor Lamme for his inspiration and central contribution to the work described in this thesis. I thank Tai-Sing Lee for stimulating collaboration. I thank Warren Slocum and Wan-Ying Chee for essential technical assistance. I thank the remaining members of the Schiller lab, past and present, for their help and camaraderie. I thank Ted Adelson, Ken Nakayama, and Earl Miller for serving on my thesis committee. Finally, I thank my family for their loving support.

Contents

Abstract.......... 2

Acknowledgments.......... 3

Introduction.......... 5

Chapter 1.......... 14
    Extra-RF modulation: cue receptivity and spatial tuning.

Chapter 2.......... 57
    Extra-RF modulation: specificity for image structure

Chapter 3.......... 88
    Extra-RF modulation: temporal properties

Summary and Conclusion.......... 99

References.......... 107

Introduction

Introduction

This thesis is about the neural representation of the world held within primary visual cortex (area V1) of the primate. Primate area V1 is an extraordinary neural structure. In the macaque monkey it consists of a vast sheet of cortex encompassing between three-hundred million and half a billion cells in each hemisphere (De Valois & De Valois, 1990). Each hemisphere codes visual information for the contralateral visual field. The visual input to area V1 comes from the lateral geniculate nucleus (LGN), but cells in this relay nucleus have essentially the same properties as their retinal ganglion cell inputs (De Valois & De Valois, 1990). Thus, in thinking about the transformation in visual processing between the eye and the cortex, we will ignore the LGN. If we consider that the number of optic nerve fibers entering the brain for one hemifield of visual space is approximately one million, we can easily see that there is a vast amplification of neural processing taking place in area V1. Numerically speaking, on average there are between three hundred and five hundred V1 neurons available to process the output of each retinal ganglion cell (Connolly & Van Essen, 1994).

Why are there so many V1 neurons available to process the retinal input? It would of course be absurd if each retinal ganglion cell merely corresponded to three-hundred or so neurons in area V1 with identical properties. In this case V1 would have accomplished nothing but replication of the retinal signal. In fact, what occurs is completely different and far more interesting--for in area V1 we find visual coding properties entirely absent from the retinal processing.

In order to see how V1 differs from the retina, we must establish some of the visual coding properties of retinal ganglion cells. The chief property of retinal ganglion cells is the concentric, antagonistic center-surround type receptive field (RF) (Kuffler, 1953; Barlow, 1953). The center-surround organization gives these cells moderate band-pass tuning for the spatial frequency of luminance contrast without any tuning for orientation (Enroth-Cugell & Robson, 1966). Thus, the retinal ganglion cells respond well to small spots of light (or dark) centered on their receptive fields, as well as to square and sine-wave grating patterns of appropriate spatial frequency (of any orientation). However, the antagonistic center-surround organization insures that the response of these cells to full-field illumination is greatly diminished. Some primate retinal ganglion cells (those projecting to the parvocellular layers of the LGN) code color; interestingly, these cells form a small number of discrete types for specific combinations of cone input (Derrington et al., 1984). Other retinal ganglion cells (those projecting to the magnocellular layers of the LGN) are not as efficient in coding color, but have greater sensitivity to luminance contrast and have much higher resolution of temporal frequency (that is, can respond to higher rates of flicker) (Kaplan & Shapley, 1982).

Before discussing how the visual coding achieved by the RFs of V1 neurons differs from the coding of retinal RFs, it is worth considering one aspect that is maintained--that is, restricted spatial scope. Although the RF size of V1 neurons is somewhat larger than that of individual retinal ganglion cells, V1 RFs are still sufficiently small to code fine-scale spatial localization (Dow et al., 1981). This is in strong contrast to the RFs of neurons in extra-striate regions, which are far larger (Desimone et al., 1985). Thus, an important point to remember as we consider the various properties of the V1 RF is that its analysis remains highly localized.

The first interesting transformation of retinal signals that was discovered in primary visual cortex is the appearance of orientation tuning (Hubel & Wiesel, 1968). Individual V1 neurons may require bar or grating stimuli of very specific orientations in order to be activated. An equally interesting transformation from retinal signals is the establishment of very narrow spatial frequency tuning (Schiller et al., 1976; Movshon et al., 1978; Albrecht et al., 1980). Thus, while retinal ganglion cells may respond to sine-wave gratings over a moderate range of spatial frequencies and of any orientation, V1 cells typically respond to a very small subset of possible sine-wave gratings--those of just the right spatial frequency and orientation.

The observation that V1 RFs can be narrowly and jointly tuned for spatial frequency and orientation has led to the proposal that these cells are engaged in decomposing visual images into their local Fourier spectrum (see De Valois & De Valois, 1990 for review). Although it remains controversial if this is exactly what these cells are doing, the localized Fourier analysis hypothesis nonetheless has remarkable predictive power. For example, when V1 neurons are presented with black and white checkerboard patterns, their responses are well predicted by the conjecture that these cells analyze separate Fourier components of the stimulus (De Valois et al., 1979). However, the conjecture that these cells code the perceptually salient black and white edges fails to predict how V1 cells will respond to the checkerboard patterns.

Another important transformation of retinal input found in the V1 neurons is the appearance of RFs with binocular input (Hubel & Wiesel, 1968). This binocular processing is thought to underlie the first stage of binocular disparity analysis, and the disparity sensitivity of V1 neurons may take several forms (Poggio et al., 1977). Furthermore, V1 neurons may display direction selectivity, a property not seen in the primate retina (Hubel & Wiesel, 1968). In addition, the coding of color becomes more complex in area V1. Instead of the small number of color coding cell types among retinal ganglion cells, V1 neurons display a wide variety of color sensitivities (Lennie et al., 1990). Interestingly, the specificity of V1 neurons to all of these visual attributes may be described with relatively simple quasilinear filters tuned in specific ways.

In the absence of information about the RF properties of V1 neurons it may have seemed baffling that there are so many V1 cells compared to retinal ganglion cells. Yet, knowing about the various characteristics of V1 RFs which indeed may come in various combinations, we can easily see how so may V1 cells could be accounted for. For example, if we imagine that different V1 RFs could be selective for different combinations of specific orientations, spatial frequencies, directions of motion, colors, and disparities, we could quickly come up with enough combinations to account for the large number of cells at each retinal locus.

We are now ready to consider an overall view of the function of V1, at least as far as the RF is concerned: V1 can be seen as a vast bank of filters. Each filter, corresponding to an RF, is tuned to a specific subset of basic visual information (such as a particular part of the local Fourier spectrum) in a localized part of the visual field. The vast array of different RF filter types and positions insures that the input visual images are split into a correspondingly large number of pieces. As the filter analogy suggests, the V1 RFs serve to separate different types of information already present in the image (e.g., different spatial frequency components, or different wavelength components). In this view of V1 RF function (which is widely, though by no means universally, held) the main contribution of V1 RFs is in *purifying* specific types of information already present in images.

The view of V1 RFs as constituting an exquisite image filtering system becomes relevant when we consider the broader task of vision. One reasonable statement of the task of vision is, given the images impinging on the retinae, the visual system must model the three-dimensional structures of the distal world. When we look out at the distal world, we have a strong experience of the actual surfaces and objects that we infer to have given rise to our retinal images. But achieving this entails far more than image filtering, which is merely the process of extracting a subset of the information already present in the input stimulus. Distal structure cannot be found through filtering, because the structures of the distal world modeled so richly in our perception do not in fact exist in the retinal images (Marr, 1982; Kanisza, 1955.) Clearly, we cannot merely extract that which does not exist. Rather, distal structure must be *inferred* from the various cues of luminance, color, and disparity that actually do exist in the retinal images (Nakayama and Shimojo, 1992). Moreover, because we have a relatively fixed vantage point of a scene at any given moment, the visual system must also make inferences about forms not directly visible, such as the manner in which surfaces complete behind occluding structures (Nakayama et al., 1989; Enns and Rensink, 1994). For the visual system to accomplish these tasks, it must have a sophisticated understanding of the way in which three-dimensional structures in the distal world can lead to the formation of retinal images.

From the discussion in the proceeding paragraph it should be clear that, to the extent that we view V1 as a bank of purifying filters, this primary visual area can play no direct role in vision's task of modeling the structures of the distal world. At best, in this view, V1 would be a preprocessing device capable of no more that transforming visual input into a convenient form for subsequent analysis. The conjecture that the function of the V1 RF is far removed from perception is in fact well supported by experiment. For example, V1 RFs are not for the most part sensitive to perceptually salient illusory contours (although neurons in area V2 do display sensitivity to these; von der Heydt, 1989). Furthermore, V1 neuron RFs appear oblivious to perceptually salient pattern motion (as, for example, with a moving plaid pattern), and respond instead to the individual components of the plaid (whereas neurons in are MT in fact display sensitivity for the pattern motion itself; Movshon et al., 1985.) Examples such as these have supported the widely held view of V1 as a visual preprocessor, with the perceptually relevant analyses of vision, the process of bringing all of the separated bits of visual information back together again, taking place in the extra-striate visual areas.

There might be little more to say about the function of area V1 in vision had the RF truly constituted the sole function of individual neurons. However, this is not the case. The RF of a visual neuron is by definition the restricted region of the visual field from which an appropriate stimulus, such as an oriented bar or a patch of texture, may drive the cell to evoke action-potential responses. Yet the activity of V1 neurons evoked in this manner may be *modulated* by stimuli placed entirely outside the RF (Maffei et al., 1976; Nelson and Frost, 1978; Gilbert and Wiesel, 1990; Knierim and Van Essen, 1992). We term this general phenomenon *extra-RF modulation.*

Extra-RF modulation is a curious phenomenon. We have discussed how the RF of the V1 neuron is localized. Yet extra-RF modulation, by enabling stimuli completely outside the RF to influence the activity of a V1 neuron, in effect renders the cell concerned with a potentially much larger region of the visual field. It is not *a priori* clear why this would be desirable. Presumably extra-RF modulation allows neurons to signal some form of comparison between the patterns inside and outside the RF (Allman et al., 1985). But the essential characteristics of extra-RF modulation, and the type of comparison that it may support, remain largely a mystery.

Although not well characterized, the modulatory influence of stimuli placed outside the RF of the V1 neuron certainly constitutes a powerful force in primary visual cortex. A dramatic demonstration of this comes from Lamme (1995), who recorded activity of V1 neurons in awake, behaving monkeys during viewing of textured displays. Lamme used textured stimuli configured such that the RF of a V1 neuron under study received an identical pattern of stimulation from trial to trial. He could then vary the pattern of texture well outside the RF. Compared to trials in which the texture was homogeneous across the entire display, Lamme found that V1 cells almost always responded more vigorously on trials in which the orientation of the texture pattern outside the RF differed from that within the RF. Lamme's experiments suggest that extra-RF modulation constitutes as robust a feature of V1 neural function as the well known RF properties of cells in this area that I described above.

The large effects of extra-RF modulation that Lamme observed bespeak the importance of this phenomenon in area V1. However, Lamme found extra-RF modulation to be of interest beyond its mere strength. Prior to Lamme, a standard view of extra-RF modulation was that it represented a relatively simple center-surround organization, analogous to that found in the retina. In this view, the normal RF formed the center mechanism, say with an orientation tuned filter. The surround (the region outside the RF) was thought to be tuned to the orthogonal orientation and to have an inhibitory influence. Thus, extra-RF modulation in this view would be a relatively simple comparison of orientation inside and outside the RF. An interesting finding by Lamme was that extra-RF modulation evoked by orientation cues could occur equally well for cells with orientation-tuned RFs and for cells whose RFs were not orientation-tuned. Thus, contrary to the simple center-surround hypothesis, which depends of RF and extra-RF function being tightly linked, Lamme's work showed that RF and extra-RF sensitivity could be dissociated. He also made an analogous observation for motion cues.

A dissociation between the function of the RF and of extra-RF modulation is interesting for a simple reason. Above we discussed the widely held view that the function of the RF is relatively simple and is relatively far removed from the interesting problems of interpreting the structure of the distal world. If extra-RF modulation's function is dissociated from RF function, the *possibility* arises that extra-RF modulation's function may be more complex and more interesting from the perspective of perception than is the function of the RF itself.

Lamme in fact chose an explanation of extra-RF modulation that granted this phenomenon a distinctively perceptual interpretation. Lamme proposed that extra-RF modulation reflects our perceptual experience of segregating figure from ground in a scene. In support of this idea he found that extra-RF modulation was evoked consistently for neurons with RFs anywhere inside of texture "figure" regions of his displays, whereas neurons with RFs over textured "ground" regions consistently responded less vigorously, even when the RF was immediately adjacent to a "figure." Coding figure versus ground in general is clearly not something that the RFs of V1 neurons could do. However, under Lamme's interpretation, the functional role of V1 neurons could be extended to the perceptual realm through extra-RF modulation.

The goal of this thesis is to advance our understanding of extra-RF modulation in area V1. The inspiration for the experiments that I describe in this thesis come from Lamme's basic idea that the function of extra-RF modulation causes V1 neurons to be more closely related to perception than could be concluded from knowledge of their RFs alone. Although only a hypothesis, this concept has proved useful in designing the experiments discussed in the following chapters.

In the first chapter I describe experiments designed to test whether extra-RF modulation is receptive to diverse cues (such as binocular disparity, luminance, and chrominance) in addition to the orientation cue already studied. The motivation for this line of research came from the simple idea that if extra-RF modulation is involved in perceptual scene segmentation (for example, of figure from ground), then extra-RF modulation should be receptive to diverse visual cues. This is because scene segmentation can be based on diverse visual cues (clearly, we can segment a figure or a surface from its background using cues other than orientation.) The results showed that extra-RF modulation is in fact receptive to these diverse cues. Also in the first chapter I examine the spatial scope of extra-RF modulation. These experiments were conducted in collaboration with Victor Lamme, Tai-Sing Lee, and Peter Schiller.

In the second chapter I describe experiments designed to further probe the functional role of extra-RF modulation. These experiments involve textured displays in which surfaces may be seen as occluded by other surfaces or structures. From the data collected in these studies I develop the hypothesis that extra-RF modulation plays a role in the perception of visual surfaces.

Finally, in the third chapter I briefly describe experiments on the temporal dynamics of extra-RF modulation in area V1. These experiments show that extra-RF modulation is a process that takes a relatively long time to appear following visual stimulation (the latency of extra-RF modulation is about double that of normal RF stimulation.) These results raise the possibility that extra-RF modulation involves feedback from higher visual areas.

Chapter 1

Extra-Receptive Field Modulation: Cue Receptivity and Spatial Tuning

Introduction

Primary visual cortex (area V1) has been the most intensely studied area of the brain; yet, a major component of its functioning remains little understood. The vast majority of neurophysiologcial research in V1 has focused on understanding the characteristics of the receptive fields (RFs) of the neurons in this area. The RF of a visual neuron is by definition the restricted region of the visual field from which an appropriate stimulus, such as an oriented bar or a patch of texture, may drive the cell to evoke action-potential responses. Yet the activity of V1 neurons evoked in this manner may be *modulated* by stimuli placed entirely outside the RF (Maffei et al., 1976; Nelson and Frost, 1978; Gilbert and Wiesel, 1990; Knierim and Van Essen, 1992). We term this general phenomenon *extra-RF modulation*. Presumably extra-RF modulation allows neurons to signal some form of comparison between the patterns inside and outside the RF (Allman et al., 1985). But the essential characteristics of extra-RF modulation, and the type of comparison that it may support, remain largely a mystery.

Although not well characterized, the modulatory influence of stimuli placed outside the RF of the V1 neuron constitutes a powerful force in primary visual cortex. A dramatic demonstration of this comes from Lamme (1995), who recorded activity of V1 neurons in awake, behaving monkeys during viewing of textured displays. Lamme used textured stimuli configured such that the RF of a V1 neuron under study received an identical pattern of stimulation from trial to trial. Compared to trials in which the texture was homogeneous across the entire display, Lamme found that V1 cells almost always responded more vigorously on trials in which the orientation, or motion, of the texture pattern outside the RF differed from that within the RF (similar to the difference between Figs. 4B and 4C).

Lamme's experiments suggest that extra-RF modulation constitutes as robust a feature of V1 neural function as the well known RF properties of cells in this area. Yet before we may integrate extra-RF modulation into a comprehensive model of the function of area V1, we must have a better understanding of the basic characteristics of this process and of the goals that it is designed to accomplish. While we have extensive information about the manner in which the RFs of V1 neurons are tuned for basic visual cues such as orientation (Hubel and Weisel, 1968; Schiller, 1976), binocular disparity (Poggio et al, 1988; Freeman et al., 1990), color and luminance (Lennie et al., 1990; Tso and Gilbert, 1988) we lack knowledge about how (or even whether) extra-RF modulation processes all of these diverse cues. While we know much about the spatial characteristics of the RFs of V1 neurons (Hubel and Weisel, 1974; Dow et al., 1981), we lack measurements of the spatial range over which extra-RF modulation may operate. And finally, while we have well-developed ideas about the functional goals that the RFs of V1 neurons may serve in localized analysis of visual information across an image (see De Valois and De Valois, 1990, for review), we lack a comparable framework for extra-RF modulation. The purpose of our two-part series on extra-RF modulation is to advance our understanding in each of these areas of uncertainty.

We conducted the experiments that we describe in this series in area V1 of awake, behaving rhesus monkeys. In this paper we focus on two questions. First, we investigate for the first time whether extra-RF modulation in individual V1 neurons may be evoked by diverse visual cues (binocular-disparity, color, and luminance) in addition to orientation already studied by Lamme. Second, we address the question, what is the spatial extent of V1 extra-RF modulation across the visual field?

Materials and Methods

Experiments were performed on three male *Macaca mulatta* weighing 8-10 kg. Two of these animals (#89-11 and #93-08) were used for extensive quantitative data collection. The third animal (#88-52), which had inferior ability to maintain fixation, was used only in initial development of this project. Prior to surgery, monkeys were trained to jump into their primate chairs and were habituated to the laboratory environment. Subsequently, each animal underwent surgical procedures for implantation of a stainless steel cranial post for restraining head position. In the same operation we implanted the given animal with a scleral search coil for monitoring eye position (Robinson, 1963). All surgical procedures were performed under deep pentobarbital anesthesia using sterile techniques; all experimental procedures were in accordance with NIH guidelines.

Following recovery from surgery, monkeys were water deprived and brought to the laboratory for training. We used a PDP-11/37 computer to regulate and monitor the monkey's behavioral tasks, to collect behavioral and neurophysiologcial data, and to signal an IBM PC for control of visual stimulation. With head restrained in the primate chair facing a computer graphics monitor, each monkey was trained to fixate small luminous spots on the screen and then to make a saccadic eye movement to a luminous target stimulus that appeared in a random position when the fixation spot was extinguished. Analog x and y eye position signals measured through the scleral search coil were collected at 200 Hz and digitized with a precision of 0.01 degrees of visual angle. For maintaining fixation and then making the correct saccades, the monkey was rewarded automatically with drops of apple juice. Supplemental juice rewards (for encouragement, as needed) were controlled manually through the computer keyboard. During training and recording, animals drank 300 to 500 ml of juice (1500 or more trials) per session. After each session animals were provided with supplementary water if they still were thirsty. Further rewards of peanuts and fresh fruit were provided once the animals returned to their home cages at the end of the day.

Stimuli were presented on an NEC multisync XL color video display unit, driven by a Number Nine Corporation graphics board with 640 x 480 pixel resolution and a frame rate of 60 Hz. The screen was 32 x 24 cm in size and was viewed at either 57 or 63 cm distance. In experiments that did not require stereoscopic stimuli, texture displays covered the entire screen. In experiments that required stereoscopic stimuli, stereo images were displayed side by side on the screen. In this case, all stimuli in each image appeared within a 9° by 9° thin white frame which remained on at all times to facilitate fusion of the stimuli. In these experiments, monkeys viewed the screen through a prism apparatus that allowed the horizontally displaced stereo images to be viewed by one eye each and to be fused at a comfortable vergence angle.

For human observers (with a separate prism apparatus for human use) disparity-defined texture stimuli produced a rich percept of a surfaces in depth. Monkey stereo-depth perception is similar to that of human beings (Poggio et al., 1977), and we presume that with appropriate presentation the display should have the same richness for monkeys. A main characteristic of binocular image fusion is that sensitivity to binocular disparity is best at the fusion depth (i.e., on the horoptor) and declines approximately symmetrically for near and far disparities (Tyler, 1983). In psychophysical tests of our monkeys' ability to detect targets defined through binocular disparity, we found exactly this pattern. Monkeys were very sensitive to a 0.05° horizontal disparity offset of a textured target from a background near the horoptor, but were increasingly less sensitive to this same offset as target and background moved to increasingly near or far disparities. This pattern of behavior would not be expected if monkeys failed to fuse the stereo images.

Monkeys initially trained to detect salient orientation-defined texture targets mastered the easier levels of the horizontal disparity task with no special training. In contrast, when targets were made visible by vertical disparity (which does not give an impression of a surface in depth but merely of binocular rivalry) monkeys did not transfer easily to this task. Monkeys also could not detect the target defined by binocular disparity when presented with a monocular image. From the combination of these results it is reasonable to believe that the monkey's perception of disparity-defined textures is similar to ours.

Neurophsyiological recording techniques: Neural recordings in awake monkeys were made through a surgically implanted cylindrical stainless steel electrode chamber (16 mm diameter) overlaying the operculum of area 17. Recording began at least three days after surgical implantation of the recording well. Microelectrodes were inserted via the oil-filled, hydraulically closed electrode chamber and through the intact dura into occipital cortex. Activity from single cells or clusters of cells was recorded extracellularly with glass coated Platinum-Iridium microelectrodes of approximately one mega-ohm impedance. The RFs of V1 neurons thus studied were in the lower contralateral visual field with eccentricities between 2° and 6°. To help insure that our microelectrodes remained in area V1, the RF positions of neurons recorded in each experiment were compared against a chart (maintained for each monkey) of the retinotopic mapping of the visual field onto striate cortex.

Within three weeks of insertion of the electrode chamber, the dura mater hardened and became covered with an epithelium up to 6 mm thick. These barriers caused difficulty with recording because microelectrodes tended to break before entering the cortex, and more importantly, because moving the microelectrode through these tissues could cause displacement of the brain. We found the latter to be highly deleterious to extra-RF modulation, perhaps because the physical displacement depressed neural activity. We took three measures to counter this problem. First, the supra-dural epithelium could be thinned through gentle aspiration (performed with the monkey under ketamine anesthesia). Second, we found that the dura became less hard if week-long breaks from recording were interspersed between weeks of recording. Third, to avoid brain displacement we moved the microelectrode through the supra-dural epithelium and the dura with the following pattern: a quick advance of about 10 micrometers, followed by a several second pause, followed by another advance, etc. In this way we avoided building pressure on the brain. The average rate of lowering the microelectrodes was approximately 1 cm per hour.

Plotting of RFs: To plot the extent of the RF of a V1 neuron under study, we moved computer-graphics-generated bars of variable size and orientation over the neighborhood of the RF as the monkey fixated. RFs boundaries were initially drawn by hand with magic-marker on an auxiliary stimulus monitor while we simultaneously watched the moving bar stimulus and monitored the evoked neural activity with an audio amplifier. Following this, we tested our estimate of RF dimensions by flashing bars of the preferred orientation inside and outside this area.

We confirmed the reliability of our RF plotting techniques by flashing texture stimuli in a region surrounding the measured RF while leaving the RF unstimulated. In Fig. 1A we show the average response of fifty single and multi-unit recording sites to direct texture stimulation of the RF. In Fig. 1B we show activity from the same sites for texture displays in which texture is excluded from a two to three degree wide region over the plotted RF region. While neurons responded vigorously to direct RF stimulation (Fig. 1A), stimulation with surrounding texture evoked at best an extremely weak response (Fig. 1B), and this typically only from multi-unit recording sites which have somewhat larger aggregate RFs than individual neurons. Thus, our RF plotting techniques were fully adequate to allow us to isolate extra-RF stimulation from direct RF stimulation.

Texture experiments: We studied each V1 neuron with static, flashed texture displays that contained the same stimulus pattern in the region over the RF from trial to trial. In some trials, the display appeared as a homogeneously textured field (e.g., Fig. 4B). In other trials, the display appeared to have a textured disc (typically 3.6° in diameter) centered on and completely covering the RF (e.g., Fig. 4A and C). Although various visual cues were used segment the disc from its background, texture within the disc was identical to that in the corresponding region of the homogeneous texture display.

We used two types of homogeneous texture display in our experiments. The first type was a true homogeneously texture display, as illustrated in Fig. 4B. More commonly we used a pseudo-homogeneous texture display constructed, for example, by pairing a standard textured disc with a background texture of the same orientation. The line terminations formed by the disc contour in the pseudo-homogeneous display served to control against the possibility that similar line terminations in other displays (e.g., in Fig. 4C) could be the source of the extra-RF modulation that we investigated. In practice, differences between these two types of texture display are only visible by careful foveal inspection. In fact, with control experiments on twenty single or multi-unit recording sites, we found that V1 neurons respond identically to the two types of homogeneous texture display. For simplicity we will ignore the distinction between the true- and the pseudo-homogeneous texture displays in the remainder of this report.

The temporal progression of a behavioral trial for most of our texture experiments described is illustrated in Fig. 2. At the beginning of a trial, a fixation spot appeared on the gray monitor screen. The monkey foveated this spot (see eye movement record at bottom of Fig. 2). Approximately 200 milliseconds after foveation of the spot occurred, a texture display appeared on the screen for a fixed interval (typically either 250 or 267 milliseconds), after which the screen returned to the pre-stimulus gray. Approximately 200 milliseconds after the texture offset, the fixation spot was extinguished and a white target spot appeared in a random position around the fixation spot. The monkey was rewarded with a drop of apple juice for maintaining stable fixation throughout the trial and then making a saccade to this target. In an alternative experimental paradigm, the monkey was required to saccade to a texture defined stimulus (either over the RF or in the opposite hemifield) following the extinguishing of the fixation spot. Operationally, stable fixation meant that the monkey's eye position remained within a control window (not visible in the stimulus display) that was centered on the fixation spot. The fixation window size varied from 1° x 1° to 0.3° x 0.3°; the typical value was 0.5° x 0.5°.

Given that the results in this study are based on comparison of neural responses in trials where the texture display was either homogeneous or contained a salient textured figure, it is of considerable importance to determine whether or not the presence of the figure in the flashed texture display could subtly influence eye movements which might in turn alter neural responses. As shown in the eye-movement record in Fig. 2, fixation was not noticeably perturbed by flashing of the texture stimulus. We addressed this topic quantitatively by selecting recordings in which neural responses showed strong extra-RF modulation depending on whether or not the texture display was of the homogeneous type or contained a texture-defined figure. For each trial, the variance in both x and y eye position was measured during the texture display interval. The pooled x and y variances are shown in histograms in Fig. 3 for both the homogeneous and non-homogeneous texture displays. The distributions are indistinguishable, and separate chi-square tests for x and y values fail to reject the null hypothesis that the content of the texture display has no influence on variance in eye position during fixation. From these results (which agree with an analysis by Lamme (1995)), we are confident that our observations of neural activity described here are not an artifact of eye movements.

Data collection and analysis: Neural spike data were collected using either hardware and software from a Brainwaves system (version 3.0) or a simple two level spike amplitude threshold discriminator. Data files containing spike and event times, and eye position signals were saved on an IBM PC in binary form and converted to ASCII for analysis on UNIX and Macintosh computer systems. Data analysis was conducted using a combination of our own C++ analysis routines and commercially available software (i.e., Mathematica and MATLAB).

Results

In this two paper series we present the results of neural recordings in area V1 in four hemispheres of two awake, behaving rhesus monkeys. Our quantitative data consists of experiments on eighty-five isolated V1 neurons and one-hundred and twenty multi-unit sites (in which inseparable signals from two or more cells were recorded simultaneously). The V1 cells that we studied had RFs in the lower, contralateral visual field with eccentricities ranging from 2° to 6° of visual angle.

Evoking extra-RF modulation

We use the expression *extra-RF modulation* to describe how a neuron's response to direct RF stimulation may be influenced by patterns appearing entirely outside the RF. The technique common to all our experiments on extra-RF modulation consists of measuring the response of a given V1 neuron to a homogeneous texture display (e.g., Fig. 4B) and using this as a standard to compare against the responses of the same cell to displays containing identical texture pattern over the RF but different patterns outside the RF area. For example, Fig. 4C shows a textured display containing a disc-shaped region that segments from the background through a 90° difference in orientation of texture elements in these two regions. In our experiments we positioned the disc so that it was centered on and completely covered the RF of a V1 neuron under study (Fig. 4A). As texture within the disc is identical to that in the corresponding region of the homogeneous texture display, stimulation with these displays allows us to isolate the effects of extra-RF modulation from the effects of RF stimulation. In the absence of extra-RF modulation, V1 neurons would respond identically to the displays in Fig. 4B and Fig. 4C. Extra-RF texture alone does not activate V1 neurons (Fig. 1).

Fig. 4D shows the responses of a V1 neuron (cell *a*) to the homogeneous texture display flashed on a gray monitor screen for 267 milliseconds as a monkey foveated the fixation spot. This Figure shows the action potentials evoked during the course of the one-hundred trials that we recorded, along with the profile of the cell's average response rate during these trials. The V1 neuron showed almost no spontaneous activity, but responded to the appearance of the homogeneous texture display with a vigorous burst of action potentials. Following this initial burst, the cell's response rate declined drastically. The cell thus maintained a much reduced level of activation for the remainder of the texture display interval, after which it returned to a quiescent state.

When we stimulated cell *a* with the orientation-defined disc (diameter 3.6°) in twenty randomly interleaved trials, we recorded dramatically different results (Fig. 4E). Although the neuron responded to the onset of the disc display with nearly the same burst of activity as in the example above, thereafter the neuron maintained a robust response level for the remainder of the disc stimulation interval. We use the gray shading in the response profile to illustrate how the response activity for the orientation-defined disc display exceeds that for the homogeneous texture display.

The data in Fig. 4 agree with Lamme's (1995) basic finding that variation in texture orientation cues well outside the RF of a V1 neuron can evoke powerful extra-RF modulation of the response to RF stimulation. This extra-RF modulation appears to be a pure sensory phenomenon. Lamme (1995) found, and we confirmed, that this modulation occurs whether or not the texture displays hold significance for the monkey's behavioral task (i.e., the effect does not depend on whether the monkey was able to ignore or had to make saccadic eye movements to the textured discs). We also found that two texture discs (one over the RF and one in the opposite visual hemifield) have the same effect as one disc over the RF, arguing against the possibility that visual attention covertly directed toward the disc display could underlie the effects we have observed (data not shown). Finally, Lamme's analysis of eye movements during texture stimulation, and our independent analysis (see Materials and Methods) indicate that it is very unlikely that small involuntary eye movements during fixation cause these extra-RF modulation effects.

Do diverse visual cues evoke extra-RF modulation?

We now turn to the first main goal of this paper: to address the question of whether other visual cues also evoke extra-RF modulation in V1 neurons. For this purpose we use the same textured disc as in Fig. 4C, but we use different cues to delineate the disc from the background texture.

Binocular disparity: We illustrate a side-view rendition of a textured disc segmented from the background through binocular disparity cues in Fig. 5A. The disc here appears to float above a textured background. As the disc texture over the RF duplicates that in the corresponding region of the homogeneous texture field, this stimulus allows us to isolate the effectiveness of disparity cues for evoking extra-RF modulation. No previous study has investigated the potential for binocular disparity cues to evoke modulation of this kind. Fig. 5B illustrates the results of stimulating cell *a* with this disparity disc display in twenty randomly interleaved trials. In this condition the neuron displayed activity similar to that for the orientation-disc display: following a burst of activity at texture onset, the cell maintained a robust rate of activity well above the response level for the homogeneous texture display. Thus, extra-RF modulation is sensitive to binocular disparity cues.

Color, luminance: In Figs. 6A and 6B we illustrate disc displays in which either color or luminance act as cues for segmenting the disc from background texture. Although previous studies have investigated effects of color on extra-RF modulation in primate extra-striate cortex (Zeki, 1973; Schein and Desimone, 1990), pure color and luminance cues have not previously been tested in this manner in primate area V1. Figs. 6C and 6D illustrate the responses of cell *a* for the color and luminance disc displays respectively. In twenty trials for each condition (randomly interleaved with those in the previous Figure), this neuron retained moderate to vigorous response rates for these disc displays compared to the homogeneous texture display. The time course of this extra-RF modulation was the same as in the previous examples.

Combination of cues: The results that we describe above show that the four individual disc-segmenting cues all evoke extra-RF modulation for our sample V1 neuron. Given this, we now consider what will happen if we use these cues in combination to segment the texture disc. Fig. 7A illustrates a rendition of the combination disc display, in which orientation, disparity, color, and luminance all serve offset the disc from the texture background. We show the activity evoked in cell *a* during stimulation with this display in Fig. 7B. In response to presentation of the combination disc display during twenty randomly interleaved trials, the cell showed a response profile typical for stimulation with the other disc displays.

In summary, with this experiment on cell *a* we have demonstrated the important new finding that extra-RF modulation in area V1 can indeed be evoked by a diverse set of cues. In the next section we address a key question raised by this finding: how prevalent are these effects of extra-RF modulation for diverse cues among cells in area V1? We proceed toward answering this question with a quantitative analysis of extra-RF modulation in our V1 sample.

Quantifying extra-RF modulation

We studied a total of sixty-five isolated V1 neurons using the textured displays described in Figs. 4 through 7, the disc in each case centered on the RF. For most cells we used discs 3.6° in diameter (n = 44). For the remaining cells we used smaller discs, though never less than 2.7° in diameter. For each cell we chose the orientation of RF texture best suited for the cell. Aside from disc position and diameter, and texture orientation, the same texture displays were used for each experiment.

Choosing a metric

In order to compare extra-RF modulation among different neurons, we require a standard metric for this phenomenon. The first problem in quantifying extra-RF modulation is to select a temporal interval over which to measure cell activity. With cell *a* we saw that consistent extra-RF modulation evolves after the cell's initial response to texture onset. In the two panels of Figs. 8 we show this to be characteristic of extra-RF modulation across the population of V1 neurons. The scatter plots in this Figure graph the response rates of each cell to the homogeneous texture display (x axis) against the average response to the disc displays (i.e., y axis = mean of response rate for orientation, disparity, color, luminance, and combination discs). In Panel A of Fig. 8 we measure the average response rates during the initial burst of activity (i.e., 50 to 100 milliseconds after texture onset). Here the points fall along the $x = y$ identity line: this indicates a lack of consistent extra-RF modulation in the initial response interval across the population of cells. In Panel B of Fig. 8 we measure the response rates during the remaining period of activation (i.e., 100 to 300 milliseconds after texture onset). In contrast to the previous Panel, here the points mostly fall above the $x = y$ identity line: these data indicate that disc displays consistently evoked greater responses than the homogeneous texture display in this later time interval. Thus, we will focus on neural activity in the interval 100 to 300 milliseconds following texture onset in our analyses of extra-RF modulation.

The second problem in quantifying extra-RF modulation involves selecting a measure that allows us to compare activity for neurons with highly disparate response rates. In Panel B of Fig. 8 we can see that the fractional difference in activity between disc and homogeneous texture displays depends little on the absolute rate of neural response. We can therefore reasonably use a measure for extra-RF modulation that factors out the absolute activity level. As our metric of extra-RF modulation evoked by a particular disc for an individual cell, we thus calculate the ratio obtained by dividing the response evoked with a given disc display by the response evoked with the homogeneous display (activity averaged over the 100 to 300 millisecond interval in all cases.) We term this metric an *extra-RF modulation ratio.*

Prevalence of extra-RF modulation for diverse cues

For every cell, we calculated these extra-RF modulation ratios comparing the responses of each disc display to the response to the homogeneous texture display (i.e., disc response / homogeneous display response). In Fig. 9 we show histograms of these modulation ratios pooled by disc type. The uppermost histogram shows data for the orientation-defined disc. In agreement with Lamme (1995), the great majority of neurons showed greater responses for the orientation-defined disc display compared to the homogeneous texture display (i.e., most values in the histogram fall above the extra-RF modulation ratio value 1.0; the median ratio value is 1.69). How do the other cues compare to the orientation case?

The next three histograms in Fig. 9 show extra-RF modulation ratios for the disparity, color, and luminance-defined disc displays. The data show that, for the great majority of neurons, each of these disc displays also evokes greater responses for these displays compared to the homogeneous display. The median extra-RF modulation ratio values for these three cues are 1.60, 1.73, and 1.33 respectively, comparable to the effect of orientation.

Extra-RF modulation ratios for the combination display appear in the bottom histogram. As with the individual disc defining cues, the V1 neurons also displayed consistently strong extra-RF modulation for the combination display in which all of these individual cues appear simultaneously. The median extra-RF modulation ratio value for this case (1.71) is comparable to previous cases. This is an interesting result because we might expect that extra-RF modulation arising from a display in which a number of potent cues segment the disc would reflect a summation of the individual effects of these cues and thus be substantially larger than extra-RF modulation evoked by any individual cue. Our data show that this is not the case. Overall, seventy-five percent of the neurons that we tested showed significantly greater responses to at least one of the disc displays compared to the homogeneous texture display ($p < 0.05$).

In summary, in this section we showed that within the population of V1 neurons, robust extra-RF modulation exists for each of the diverse cues that we tested. These results are important because they suggest that extra-RF modulation may serve a function that generalizes across visual cues. If widespread extra-RF modulation had instead existed for only a subset of the disc displays (say, those defined by orientation and luminance, but not those defined by color or disparity), this phenomenon could at best serve only a restricted role tied to particular visual cues.

Comparing effects on individual cells of extra-RF modulation from diverse cues

Range of extra-RF cue independence and selectivity

We now turn to a question not answered by the population analysis above: to what extent does extra-RF modulation generalize across cues for individual neurons? We saw above that the effects of extra-RF modulation manifest themselves through the maintenance of greater cell responses for all disc displays compared to the homogeneous display. Yet depending on the relative strength of responses for different disc displays, the nature of the information conveyed through extra-RF modulation could assume very different characters in individual neurons.

At one extreme, a neuron that responded with equal strength to each of the disc displays would show a generalization of extra-RF modulation across cues. We show data from such a neuron (cell *b*) in Fig. 10A, which illustrates the response of this cell to the homogeneous texture display and to each of the five disc displays. The nearly identical large responses for each of the disc displays compared to the homogeneous texture display indicates that extra-RF modulation in this neuron occurs largely independent of the identity of the visual cues that evoke it.

One simple explanation for the invariance in response to disc displays that we see with cell *b* is that the cell reaches some saturating level of activation that causes the response to each disc display to converge to the same activity level. We can counter this argument by simply showing that the neuron in fact did *not* reach saturating levels of activity during stimulation with the normal texture displays. We found that with a completely different RF stimulus, we were able to elicit a response rate much larger than that evoked by any disc display. For cell *b*, stimulation with texture in the right eye only was a very powerful stimulus. As seen in the response profile in Fig. 10B, the activity evoked by this monocular stimulation in randomly interleaved trials far exceeds activity evoked by the normal disc displays, demonstrating that the disc displays did not in fact saturate the cell's response. Thus, the cue-invariance of extra-RF modulation for this cell must arise not from some trivial saturation in neural response, but from an intrinsic property of extra-RF modulation itself.

For descriptive purposes, we define cue-invariance of extra-RF modulation to mean that the difference in response activity between a cell's best and least preferred textured discs is smaller than the difference in response activity between the least preferred disc and the response to the homogeneous texture display. Under this definition, approximately one-eighth of the cells in the sample are cue invariant (i.e., 8 / 66). Fig. 11A illustrates representative examples of four of these cells (including cells *a* and *b*). Each neuron is represented by the extra-RF modulation ratios for each of the five disc displays, ranked in ascending order of strength.

In contrast to these clear examples of cells with cue-invariant extra-RF modulation, we did not find as compelling evidence of *cue-selective* modulation extra-RF modulation. Fig. 11B illustrates one of the clearest examples of extra-RF cue selectivity in our sample (cell *c*). Despite a strong bias for one extra-RF cue (in this case, binocular disparity), this cell was not purely disparity selective, as the color-defined disc (the second strongest) also evoked modest extra-RF modulation.

In Fig. 11C we illustrate a cell with intermediate characteristics. Although not completely balanced in sensitivity to different cues, the cell clearly showed extra-RF modulation for each of the disc displays. The remaining cells in our sample form a continuum between the cue-invariance and cue-selectivity of the examples in Figs. 11A and B. As we will see in the next section, this insures that the great majority of cells receive convergent information from the extra-RF visual cues that we tested.

Correlated effect of extra-RF modulation for pairs of cues

As a quantitative approach to examining effects of diverse cues on individual neurons, we compare the efficacy of various pairs of disc-defining cues for eliciting extra-RF modulation. In Fig. 12A we illustrate a scatter plot comparing extra-RF modulation ratios for orientation-defined discs versus those for disparity-defined discs for each cell in our sample. If most neurons had extra-RF modulation selective for either orientation or disparity, the points in this graph should fall either along the line $x = 1$ or the line $y = 1$. Instead, the points representing the cells in our sample are correlated and mostly fall along the diagonal line $x = y$. It follows that a V1 neuron which receives extra-RF modulation from a disc defined by orientation cues will likely receive a similar degree of modulation from a disc defined by disparity cues. Cells *a* and *b* fall directly on the identity line, whereas cell *c*, one of the most extra-RF cue selective cells in our sample, appears as an outlier. The non-parametric Spearman correlation coefficient (r_S) provides us with a quantitative description of the relationship between extra-RF modulation for the two disc defining cues. The value of r_S in this case equals 0.74 .

With our five different disc displays, we can make ten pairwise comparisons of the type illustrated in Fig. 12A. We list correlation coefficients for each comparison in Table 1. In each case the relationship between pairs of disc defining cues was qualitatively similar to the orientation versus disparity case, although r_S values were somewhat lower. Extra-RF modulation for luminance and color-defined discs was the least well correlated ($r_S = 0.38$), although in this case (as with all others), the correlation was highly significant ($p < 0.01$).

Despite the convergence of extra-RF signals described above, we have also seen that these signals do not seem to summate within the context of the combination display (recall cells *a* and *b* and the population analysis in Fig. 9). We further examine this issue in Fig. 12B, a scatter plot of extra-RF modulation ratios for the combination display (*x* axis) versus the mean of the extra-RF modulation ratios from each of the disc displays defined by individual cues (*y* axis). The points in this plot cluster along the identity line $x = y$. Here then is a demonstration at the individual cell level of the approximate equivalence of the combination display with each of the individual disc displays.

In summary, the data in Fig. 12 and in Table 1 present a picture of enormous convergence of extra-RF signals from diverse cues onto individual V1 neurons. The data do not indicate the existence of major V1 neural subsystems devoted specifically to extra-RF analysis of individual visual cues (i.e., parallel pathways for specific cue information). Rather, the correlation of extra-RF modulation ratios that we measured for each of the pairs of cues suggests a continuity of processing generalized across these diverse types of visual input.

"Disc alone" display

Up to this point we have focused on disc displays in which the disc segments from a high contrast textured background through either a specific visual cue or a combination of cues. However, perhaps the most obvious way to visualize the texture disc is to have no background texture at all. In Fig. 13A we illustrate a display of this type, called the "disc alone" condition. The texture disc in this case is identical to that in other displays. We tested this display with forty-four of the V1 neurons described above, interleaving this display randomly among trials with other displays. In trials in which the "disc alone" condition appeared, the area around the disc remained an unchanged gray.

In Fig. 13B we plot extra-RF modulation ratios for the orientation-defined disc (x axis) against extra-RF modulation ratios for the "disc alone" condition (y axis) for each cell thus tested. The data points are well correlated along the identity line (r_S = 0.70). These results in Fig. 13B show that defining the disc by texture orientation cues and by simply having no background texture evokes very similar extra-RF modulation in V1 neurons.

Spatial extent of extra-RF modulation

We measured the spatial extent of extra-RF modulation by varying disc diameter (while keeping the RF centered). We studied fifty-three single or multi-unit sites in these experiments. For most experiments we used the cue that evoked the most powerful extra-RF modulation at the given recording site. We used only orientation, color, or luminance cues for this part of the study, so that the entire monitor screen could be covered with binocularly viewed texture.

Fig. 14 illustrates the mean modulation of these recording sites for a range of disc diameters from 1.8° to 14.4° . The magnitude of extra-RF modulation falls monotoncially with disc diameter, becoming indistinguishable with the full screen homogeneous texture display (32° x 24° in size) at approximately 10° diameter. The data are well fit by the function

$$\text{disc response}(\ d\ )\ /\ \text{homogeneous response} - 1 = (\ d\ /\ 2\ )^{-1.7}$$

where d is the disc diameter. This result indicates that extra-RF modulation is approximately inversely proportional to disc area. This smooth, monotonically falling spatial tuning function was typical of data at individual recording sites in our sample. Only at the smallest disc diameter (1.8°) did we occasionally find deviations from this pattern (perhaps reflecting an interaction between the disc contour and the RFs of neurons in these cases.) From the smooth spatial tuning function it seems that extra-RF modulation involves continuous portions of the visual field outside the RF (at least in the radial dimension) rather than either discrete radial bands or "hot spots."

On the graph in Fig. 14 we mark the range of disc diameters used in the multiple-cue experiments described above (i.e., 2.7° to 3.6°) for comparison (vertical dotted lines). On the one hand, we observed extra-RF modulation for disc diameters substantially larger than this range of diameters. On the other hand, discs of smaller diameter than these evoked even larger strength extra-RF modulation. The spatial extent of the extra-RF modulation that we measured is greater than anything previously demonstrated in primate V1 through neural recording.

Discussion

In this paper we investigated basic characteristics of extra-RF modulation among neurons in area V1 of the awake, behaving primate. We use the expression *extra-RF modulation* to describe how a neuron's response to direct RF stimulation may be influenced by patterns appearing entirely outside the RF. We use the term *modulation* in describing this effect because, although extra-RF stimuli do influence neural activity, they do not alone directly drive neural responses without RF stimulation, as is easily verified (as with the control in Fig. 1). In our experiments we observed extra-RF modulation by comparing the responses of individual cells to homogeneously textured displays with responses to other displays that had identical texture patterns in a region covering the RF but different texture outside this region. Behavioral controls and analysis of eye movements by Lamme (1995), which we replicated in this study, indicate that the extra-RF modulation that we describe in V1 is a pure sensory phenomenon.

Characteristics of V1 extra-RF modulation

In the first part of this paper we ask whether diverse cues may serve to evoke extra-RF modulation in V1 neurons. Of the set of cues that we tested for each cell (i.e., orientation, binocular-disparity, color, and luminance), only orientation had previously been shown to effect extra-RF modulation in primate V1 (Lamme, 1995; Knierim and Van Essen, 1991). Thus, it was certainly conceivable before beginning our experiments that the remaining cues might either not evoke extra-RF modulation at all, or that they might do so in a distinctly different manner for each cue. Instead, we found that V1 neurons do maintain sensitivity for binocular-disparity, color, and luminance cues outside the RF. The extra-RF modulation for these diverse cues was of the same nature as for orientation cues, in that extra-RF modulation for these textured disc displays consistently resulted in responses greater than for the homogeneous texture display. As Lamme found with the orientation cue (1995), the onset of extra-RF modulation for the other cues occurred with a delay relative to the initial visual responses of V1 neurons. Strong extra-

RF modulation was consistently evolved by one-hundred milliseconds after the texture stimulus onset. Seventy-five percent of cells in our sample displayed significant extra-RF modulation for at least one of the texture disc diplays.

We found that in some neurons, the modulation evoked by all four diverse extra-RF cues was similar in magnitude (e.g., cells *a* and *b*). Remarkably, this property of cue-invariance even held for these cells when discs segmented from their background texture by a combination of all four of these diverse cues simultaneously. Approximately one-eighth of cells in the sample displayed this form of cue invariant extra-RF modulation. The remaining cells showed extra-RF cue sensitivity intermediate between invariance and a moderate selectivity for particular cues (the latter best exemplified by cell *c*). As the population analyses in Figs. 9 and 12 and in Table 1 show, convergence of diverse cues onto individual V1 neurons is a dominant characteristic of extra-RF modulation. This means that cues traditionally considered separate subjects of study, such as color and binocular-disparity, are linked in the sense that extra-RF modulation commonly uses them both. Although it has been suggested that different visual cues, such as color and binocular disparity, are processed independently by separate anatomical modules in the visual system (Livingstone and Hubel, 1988), our results show that many V1 neurons treat these cues interchangeably, at least from the perspective of extra-RF modulation.

Lamme (1995) previously found that motion cues can also evoke extra-RF modulation in V1 neurons. We did not test motion cues in our experiments, but it seems to us very likely that the cue-invariance that we observed in some V1 neurons would extend to the motion condition as well.

In the second part of this paper we provided the first measure of the spatial tuning of extra-RF modulation in area V1 through recordings of individual or clusters of neurons. These data showed extra-RF modulation to operate over a far greater range than previously known (approximately 10° diameter on average for cells with eccentricity between 2° and 6° of visual angle). Taken together, our results support a novel view of V1 in which individual neurons co-process information from diverse cues over a spatial extent very large compared to their RFs.

RF versus extra-RF characteristics in V1

In the past, studies of the influence of stimuli outside the RF have been conducted from the perspective that the function of extra-RF modulation is closely tied to the properties of the RF itself. For example, Gilbert and Wiesel (1990) showed in cat striate cortex that oriented bar stimuli outside the RF may influence the RF orientation tuning of neurons in this area. One interpretation of this result is that extra-RF modulation serves to put a normal RF function into a broader spatial context, but does not qualitatively alter the role of V1 neurons as devices for signaling orientation information.

In contrast, work by Lamme (1995) counters the notion that extra-RF modulation and RF function must be tightly linked in V1 neurons. For example, Lamme found that extra-RF modulation may be evoked by orientation or motion cues outside the RF even in V1 neurons that lacked RF tuning for these cues. In fact, Lamme found no correlation between the selectivity of V1 neurons' RFs for these cues and the strength of extra-RF modulation evoked by the same cues placed well outside the RF. These data cannot be explained simply by assuming that RFs of neurons under study were larger than the experimenter believed, because the data show that the properties of extra-RF modulation and the RF tuning need not be closely related.

In the present study we examined the RF tuning characteristics of some V1 neurons, such as cell *a*. The RF of this cell lacked tuning for binocular disparity or orientation, but did display color selectivity. Nonetheless, cell *a* showed strong extra-RF modulation from each of these visual cues presented well outside its RF. Similar tests on other cells, combined with Lamme's work, suggest to us a strong dissociation between the properties of V1 neurons' RFs and the properties of extra-RF modulation.

The characteristic temporal delay in the onset of extra-RF modulation may provide an opportunity for separating a V1 neuron's signals pertaining to the RF alone from those that mix RF and extra-

RF signals. In contrast to extra-RF modulation, the tuning properties of the RFs of V1 neurons for cues such as orientation are consistently evident from the first action potentials elicited at the onset of a stimulus (Celebrini et al., 1993).

Possible mechanisms underlying extra-RF modulation

Excitation versus *release from inhibition*

The extra-RF modulation that we observed was virtually always manifested as a greater response to the various disc displays compared to the homogeneous texture display. This difference in activity could arise in a number of ways. For example, the greater activity for disc displays could come from *excitatory* signals arising outside the RF. There are, however, two problems with this hypothesis. First, in the "disc alone" condition we observed responses typical for the other disc displays, although there were in this case no texture stimuli outside the disc that could have given rise to excitatory extra-RF signals. Second, the idea that excitatory signals arising from outside the RF underlie extra-RF modulation is problematic because such signals would tend to confound the V1 neurons signals about the pattern in the RF itself.

An alternative explanation is that the difference in response between the homogeneous texture display and the disc displays results from a *release from inhibition* for the latter condition. In the case of the "disc alone" condition, the very lack of texture outside the disc might be the cause of a release from inhibition. Furthermore, the release from inhibition hypothesis would mean that when a V1 cell is more active, it conveys more faithful signals about the local pattern within its RF. This seems like a more sensible situation than the reverse, which would hold with the hypothesis in the previous paragraph. As an example, consider the oscillations in the response profile of cell *a*. These oscillations are at 60 Hz and almost certainly reflect the ability of the neuron's RF to follow the vertical refresh rate of our graphics monitor. The oscillations are far more prominent for the disc displays than for the homogeneous texture display. It seems to us more plausible that this would arise from a release from inhibition of the cell's RF-driven activity than from the introduction of extra-RF excitatory cortical signals with timing precise enough to cause these prominent oscillations.

The data that we present in Fig. 8B, which show that the fractional difference in activity between disc and homogeneous texture displays depends little on the absolute rate of neural response, is consistent with idea that the extra-RF modulation acts through a release from divisive inhibition. Divisive intracortical inhibition has previously been reported as a physiological mechanism in cat striate cortex (Bonds, 1989) and has been suggested to underlie intracortical feedback signals with a temporal delay compatible with extra-RF modulation (Wilson and Humanski, 1993).

Lateral and feedback connections

There are multitudinous anatomical pathways by which the signals underlying extra-RF modulation could converge onto individual V1 neurons. The temporal delay before extra-RF modulation is consistently evolved (up to one-hundred milliseconds after the stimulus onset) would permit the visual signals to travel great distances within the visual brain before converging on the V1 neurons we have studied. In area V1, lateral connections consisting of serial chains of many interneurons, or longer range connections in superficial cortical layers (Rockland and Lund, 1983), could both be involved in the convergence of signals from outside the RF. Furthermore, feedback connections from extra-striate cortical areas (where RFs are larger in spatial extent; Desimone et al, 1985) could also contribute to extra-RF modulation. Our experiments up to this point do not allow us to distinguish among these various possibilities.

Figure Legends

1 Isolating the RF. **A** shows the average response profile of fifty single or multi-unit recording sites to direct texture stimulation of the RF. **B** shows activity of the same sites to texture stimulation when texture is excluded from a two to three degree wide region over the plotted RF region. The general lack of response shows that our RF plotting techniques are adequate to isolate extra-RF stimulation from direct RF stimulation.

2 Temporal progression of behavioral trial. All elements of the Figure are displayed relative to the time scale marked on the *x*-axis. The top three traces illustrate the time course of the fixation spot, the saccade target, and the texture display. Below is illustrated an eye movement trace collected during one trial.

3 Eye position variance for texture presentation. The **top panel** is a histogram of the standard deviation of eye position during homogeneous texture display presentation during several hundred trials. Measures for *x* and *y* eye positions were measured separately and pooled. The **bottom panel** shows data in the same format for non-homogeneous texture displays (i.e., displays with some figure element such as a textured disc). The distributions are indistinguishable, and separate chi-square tests for *x* and *y* values fail to reject the null hypothesis that the content of the texture display has no influence on the variance in eye position during fixation.

4 Extra-RF modulation with orientation cues. **A** illustrates the configuration of stimulus presentation. **B** illustrates a portion of a homogeneous texture display. **C** illustrates an orientation-defined disc. **D** illustrates the response of a V1 single unit (cell *a*) to the homogeneous texture display. **E** illustrates the response of the same cell to the orientation-defined disc. The response to the homogeneous display is shown superimposed for comparison. The gray shading illustrates where the response to the disc display exceeds that to the homogeneous texture display.

5 Extra-RF modulation with binocular disparity cues. **A** illustrates a schematic of a disparity-defined disc. Shadow and perspective cues were not in the actual display. **B** illustrates the response of cell *a* to this display.

6 Extra-RF modulation with color or luminance cues. **A** illustrates a color-defined disc. **B** illustrates a luminance-defined disc. The response of cell *a* to the color defined disc is shown in **C**. The response of cell *a* to the luminance defined disc is shown in **D**.

7 Extra-RF modulation with a combination of cues. **A** illustrates a schematic of a disc defined by orientation, disparity, color, and luminance cues. Shadow and perspective cues were not in the actual display. **B** illustrates the response of cell *a* to this display.

8 Extra-RF modulation in early and late phases of response. Both panels of this Figure are scatter-plots comparing response to homogeneous texture displays (in spikes/s) with the average response to the five disc displays (i.e., average response to orientation, disparity, color, luminance, and combination discs, in spikes/s). Sixty-five isolated cells are represented. In **A** the measure of activity is over the interval 50 to 100 ms after texture onset. In **B** the measure of activity is over the interval 100 to 300 ms after texture onset.

9 Extra-RF modulation ratio histograms for diverse cues. Each histogram compiles extra-RF modulation ratios for sixty-five isolated cells. Histograms for each of the five disc types are shown separately.

10 Cue-invariance not tied to mere saturation of cell response. **A** shows response profiles for a single V1 neuron (cell *b*) to the homogeneous texture display and the various disc displays. Extra-RF modulation in this cell appears cue-invariant. **B** shows the response of the same cell to completely different RF stimulation (monocular

texture). The response to this preferred stimulus is far greater than to any of the responses in **A**, showing that the cell had not reached response saturation when it displayed cue-invariant behavior.

11 Examples of extra-RF modulation in different V1 neurons. **A** shows responses of cue-invariant cells. Each cell is represented by a bar graph where the height of the bar represents neural response rate. The left-most bar with cross-hatching represents the response to the homogeneous texture display. The five black bars represent to the responses to the five different disc displays, ranked in ascending order of efficacy for evoking extra-RF modulation. **B** shows a cue invariant cell in the same format. **C** shows a cell with intermediate properties.

12 Scatter plots of extra-RF modulation ratios. **A** shows extra-RF modulation ratios for orientation-defined discs against those for disparity-defined discs. **B** shows extra-RF modulation ratios for the combination disc against the mean (across each cell) extra-RF modulation ratios for discs defined by individual cues.

13 Extra-RF modulation for the "disc alone" display. Above is shown an example of the "disc alone" display. Below is a scatter plot of extra-RF modulation ratios for the orientation-defined disc against those for the "disc alone" display.

14 Spatial extent of extra-RF modulation. The top panel shown extra-RF modulation ratio as a function of disc radius. Data are averaged over fifty-three single or multi-unit recording sites. The lower graphs shows that the data are well described by a power function with exponent of -1.7.

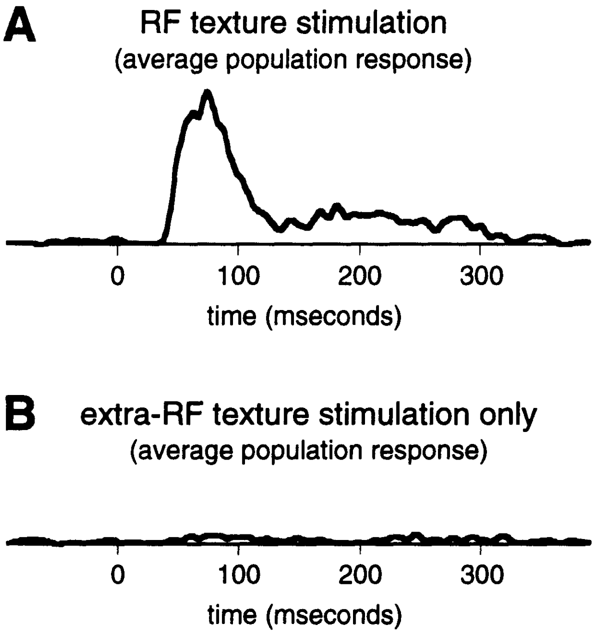

A
RF texture stimulation
(average population response)
0
100
200
300
time (mseconds)
B
extra-RF texture stimulation only
(average population response)
0
100
200
300
time (mseconds)

Figure 1

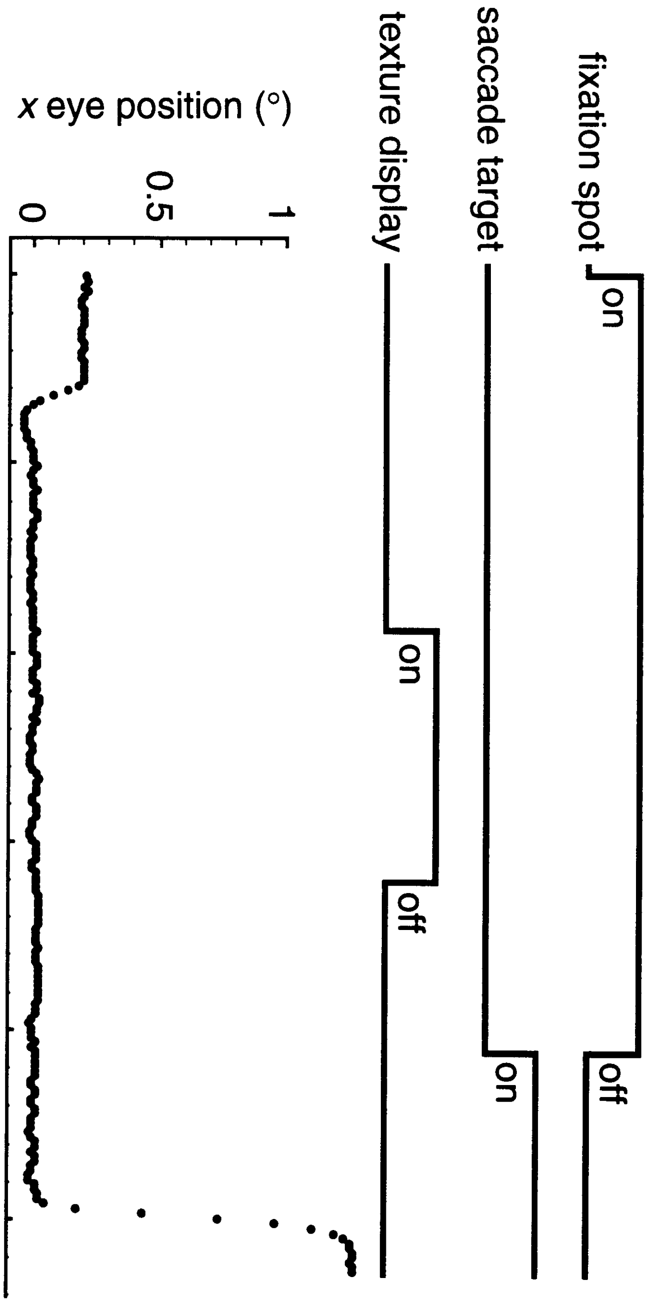

fixation spot
on
off
saccade target
on
texture display
on
off
x eye position (°)
1
0.5
0
0
200
400
600
800
1000
time (mseconds)

Figure 2

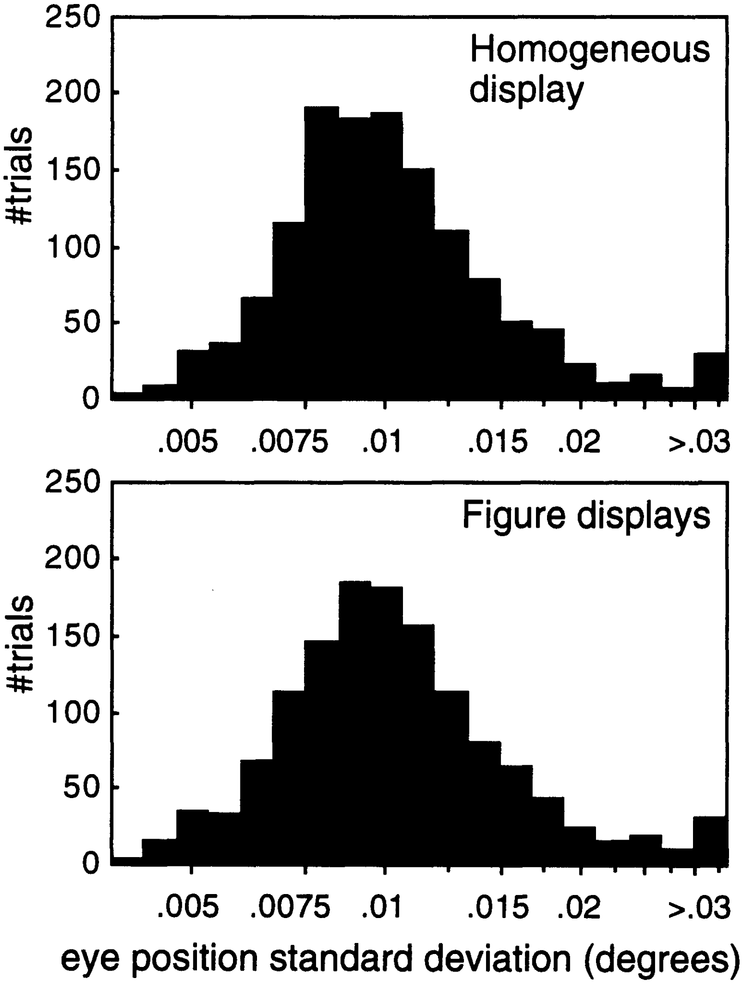

250
200
150
100
50
0
#trials
Homogeneous display
.005
.0075
.01
.015
.02
>.03
250
200
150
100
50
0
#trials
Figure displays
.005
.0075
.01
.015
.02
>.03
eye position standard deviation (degrees)

Figure 3

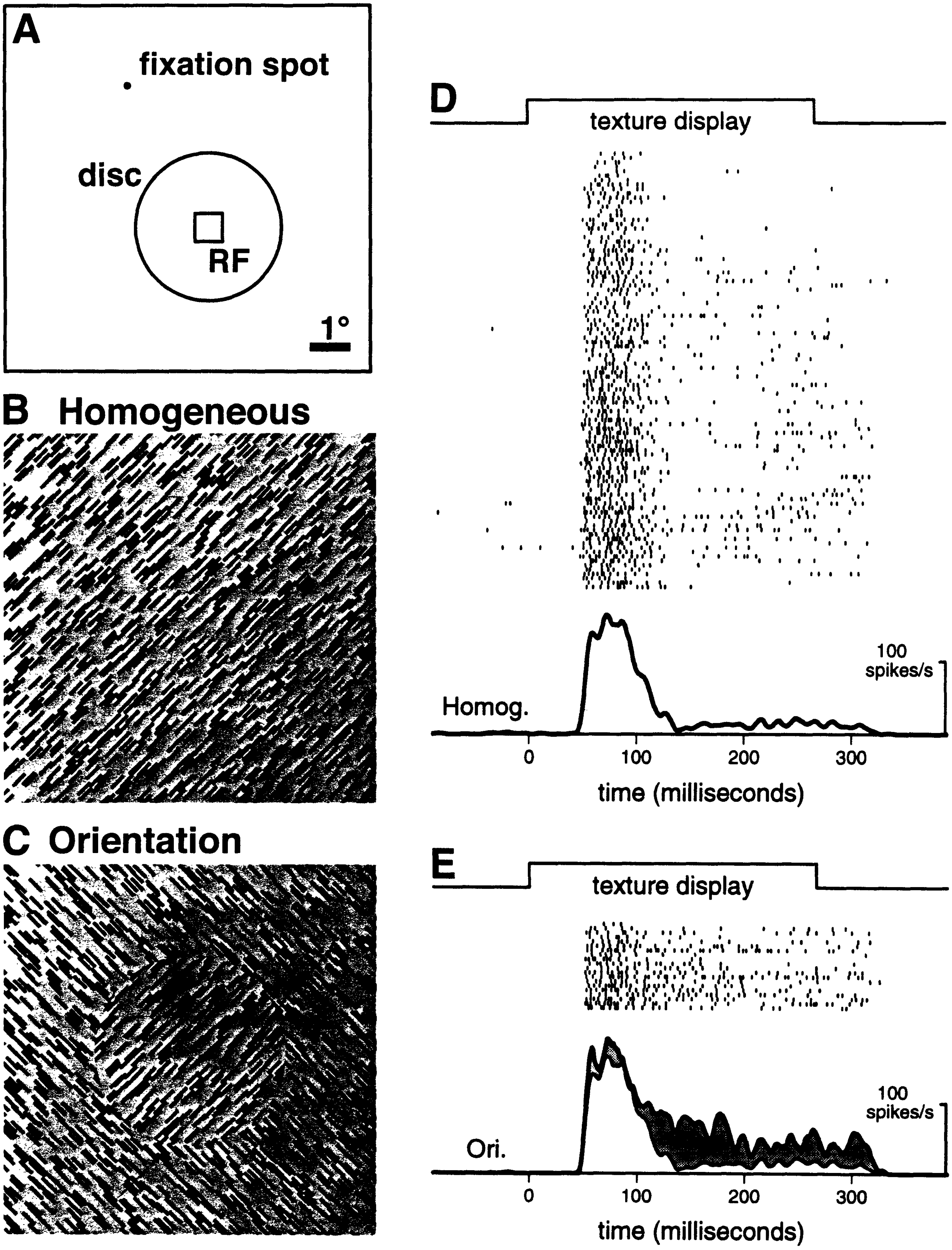

A
fixation spot
disc
RF
1°
B Homogeneous
C Orientation
D
texture display
Homog.
100
spikes/s
0
100
200
300
time (milliseconds)
E
texture display
Ori.
100
spikes/s
0
100
200
300
time (milliseconds)

Figure 4

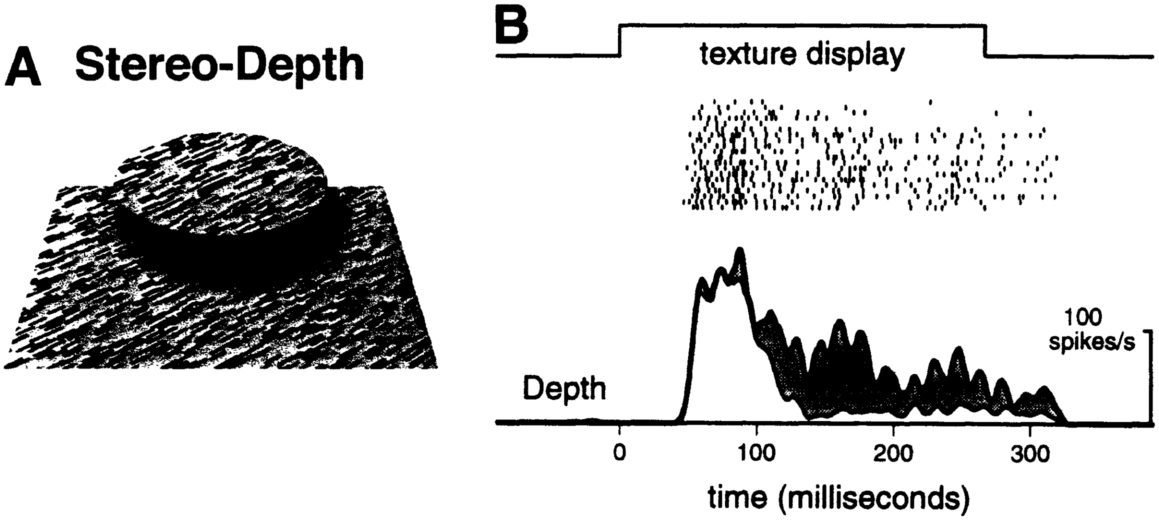

A Stereo-Depth
B
texture display
100
spikes/s
Depth
0
100
200
300
time (milliseconds)

Figure 5

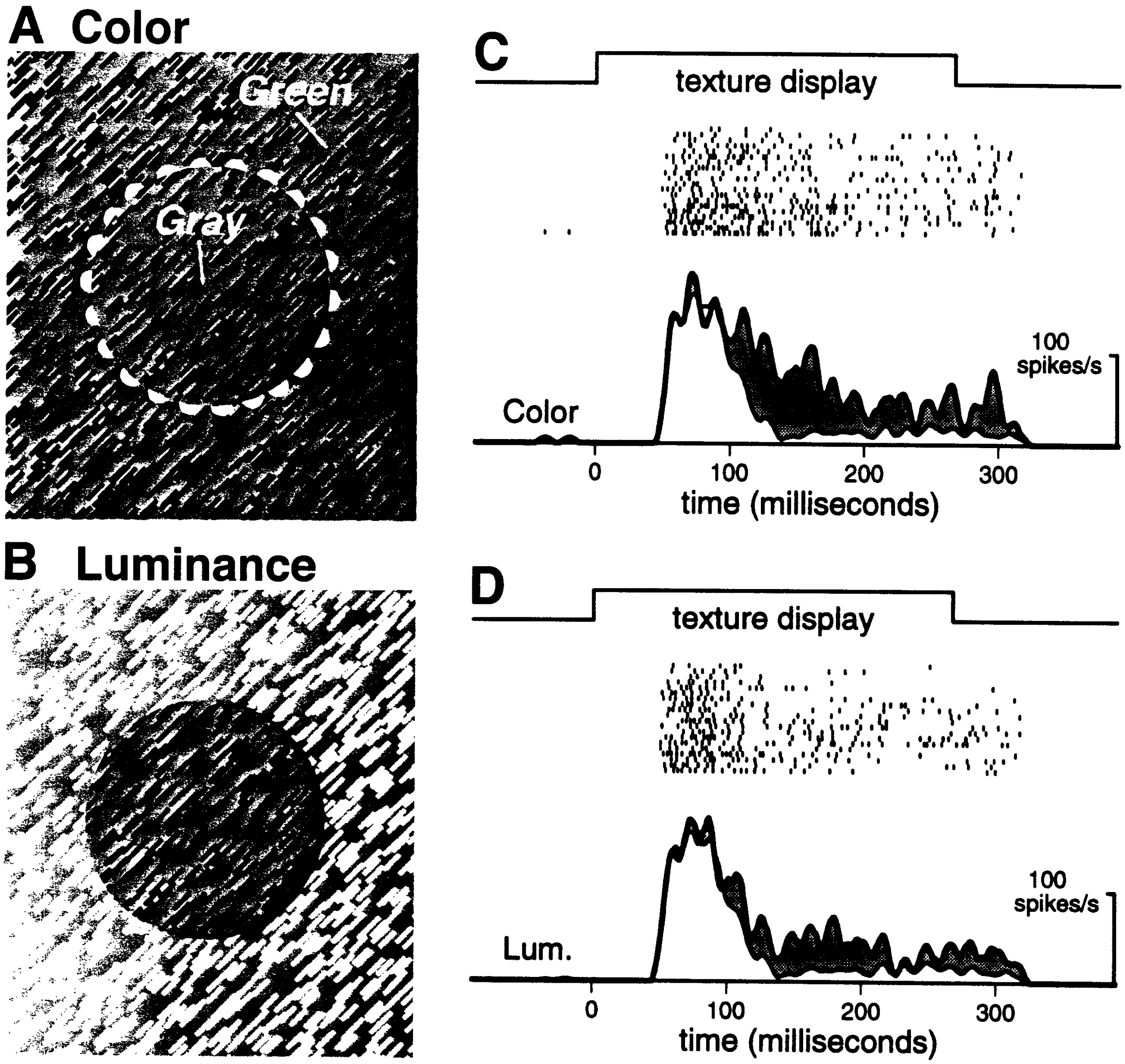

A Color
Green
Gray
B Luminance
C
texture display
Color
100
spikes/s
0
100
200
300
time (milliseconds)
D
texture display
Lum.
100
spikes/s
0
100
200
300
time (milliseconds)

Figure 6

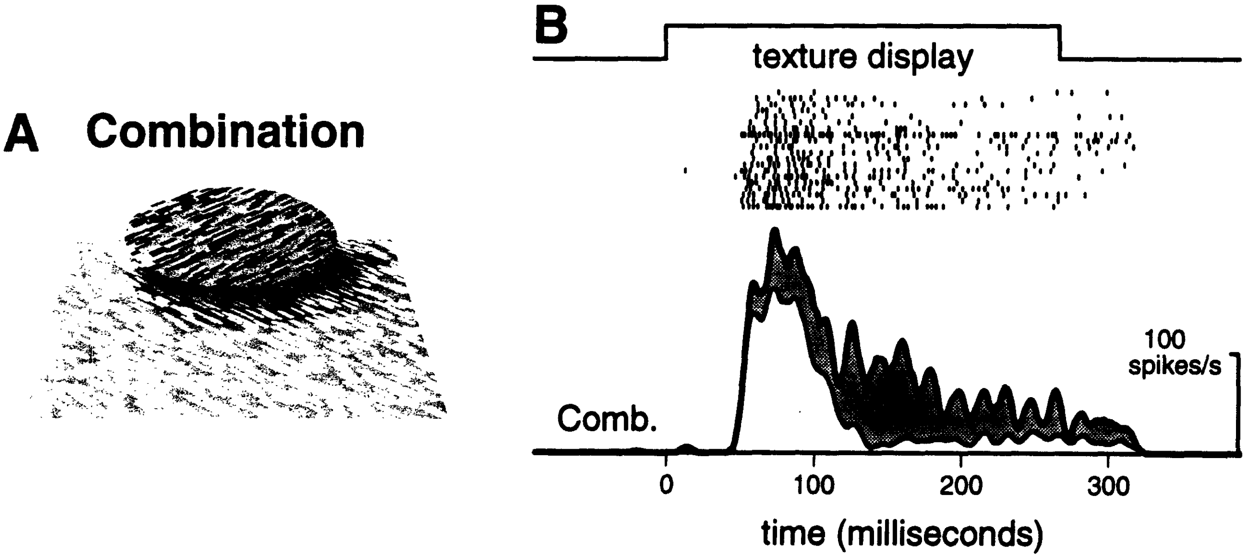
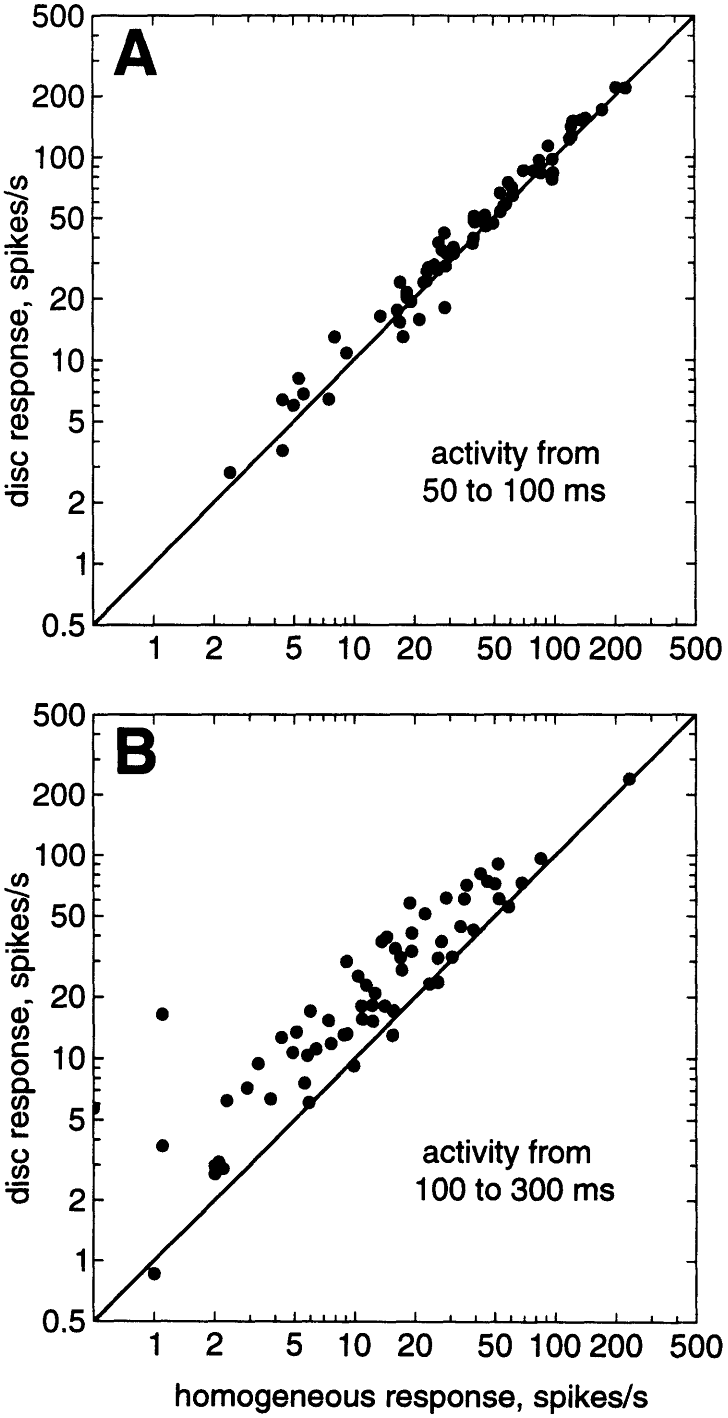

A
Combination
B
texture display
Comb.
100
spikes/s
0
100
200
300
time (milliseconds)

Figure 7



A
B
disc response, spikes/s
homogeneous response, spikes/s
500
200
100
50
20
10
5
2
1
0.5
1
2
5
10
20
50
100
200
500
activity from
50 to 100 ms
activity from
100 to 300 ms

Figure 8

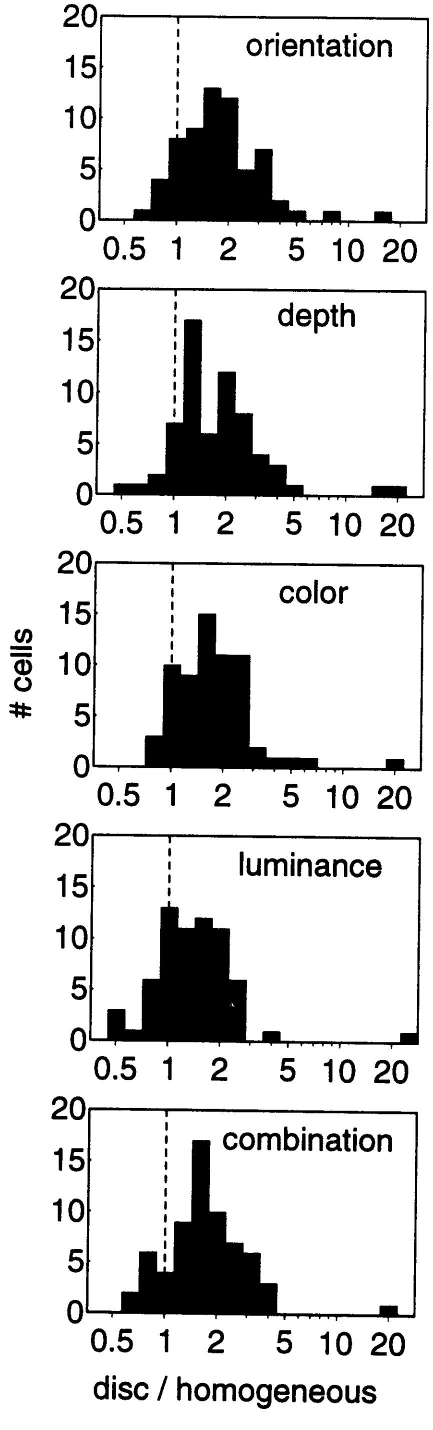

orientation
depth
color
luminance
combination
20
15
10
5
0
0.5
1
2
5
10
20
cells
disc / homogeneous

Figure 9

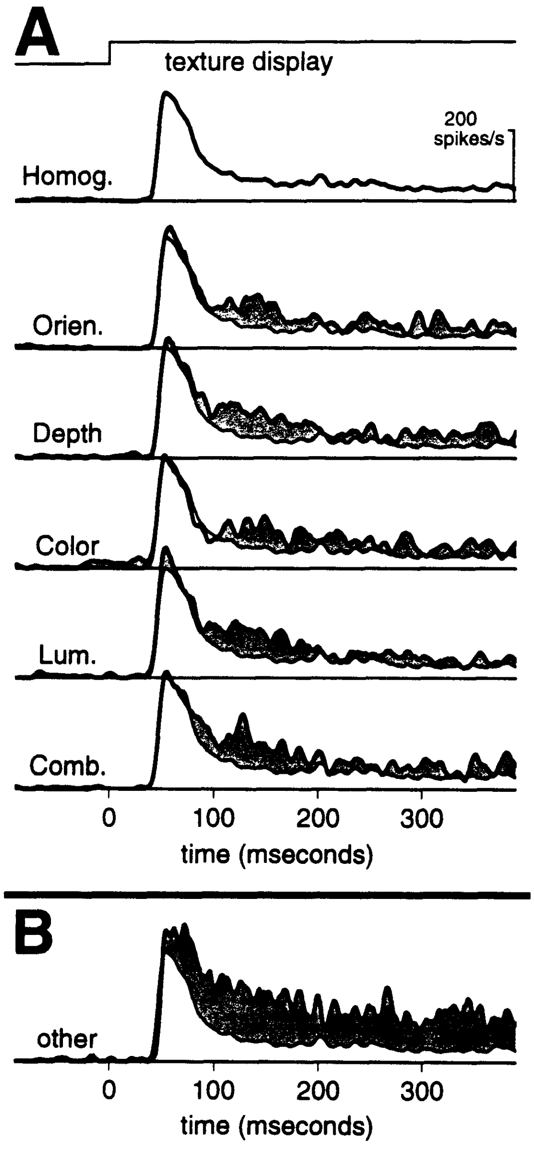

A
texture display
200
spikes/s
Homog.
Orien.
Depth
Color
Lum.
Comb.
0
100
200
300
time (mseconds)
B
other
0
100
200
300
time (mseconds)

Figure 10

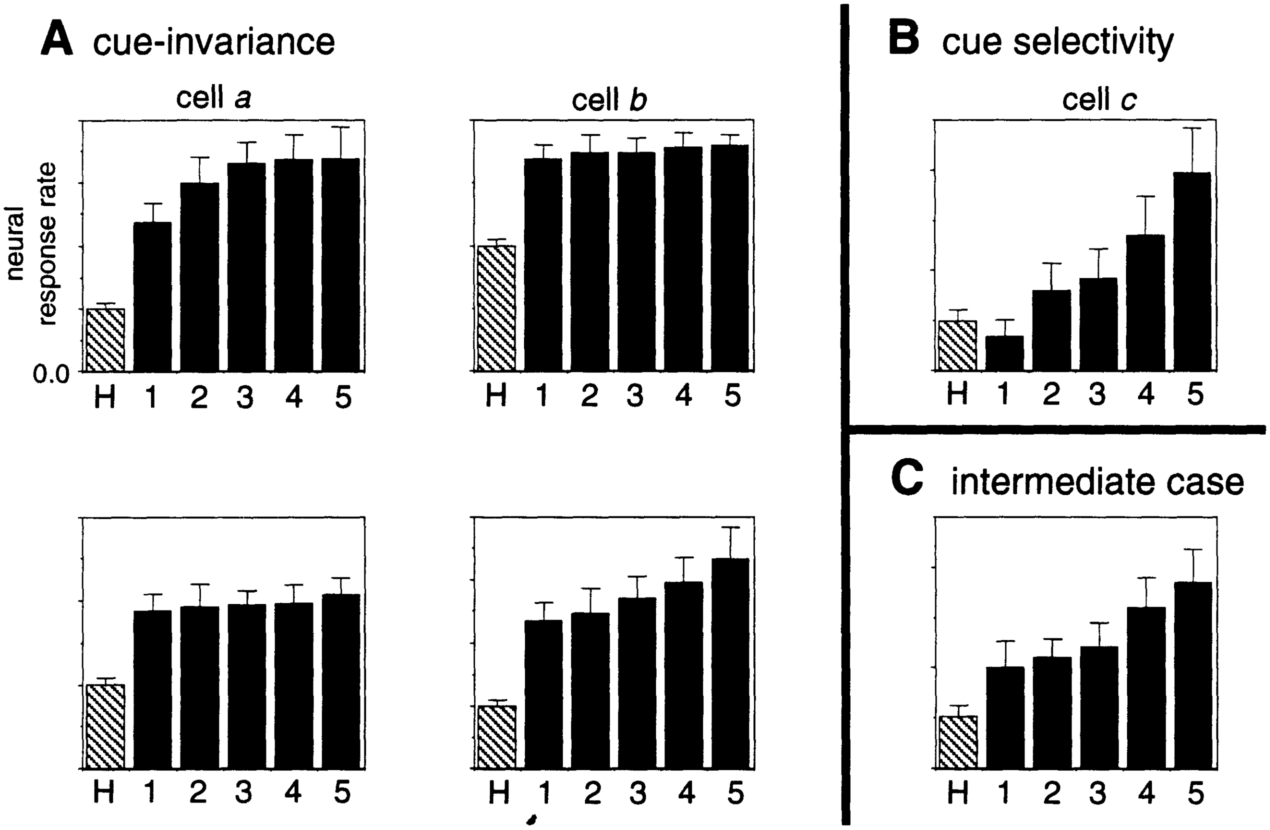

A cue-invariance
cell a
cell b
neural response rate
0.0
H 1 2 3 4 5
H 1 2 3 4 5
H 1 2 3 4 5
H 1 2 3 4 5
B cue selectivity
cell c
H 1 2 3 4 5
C intermediate case
H 1 2 3 4 5

Figure 11

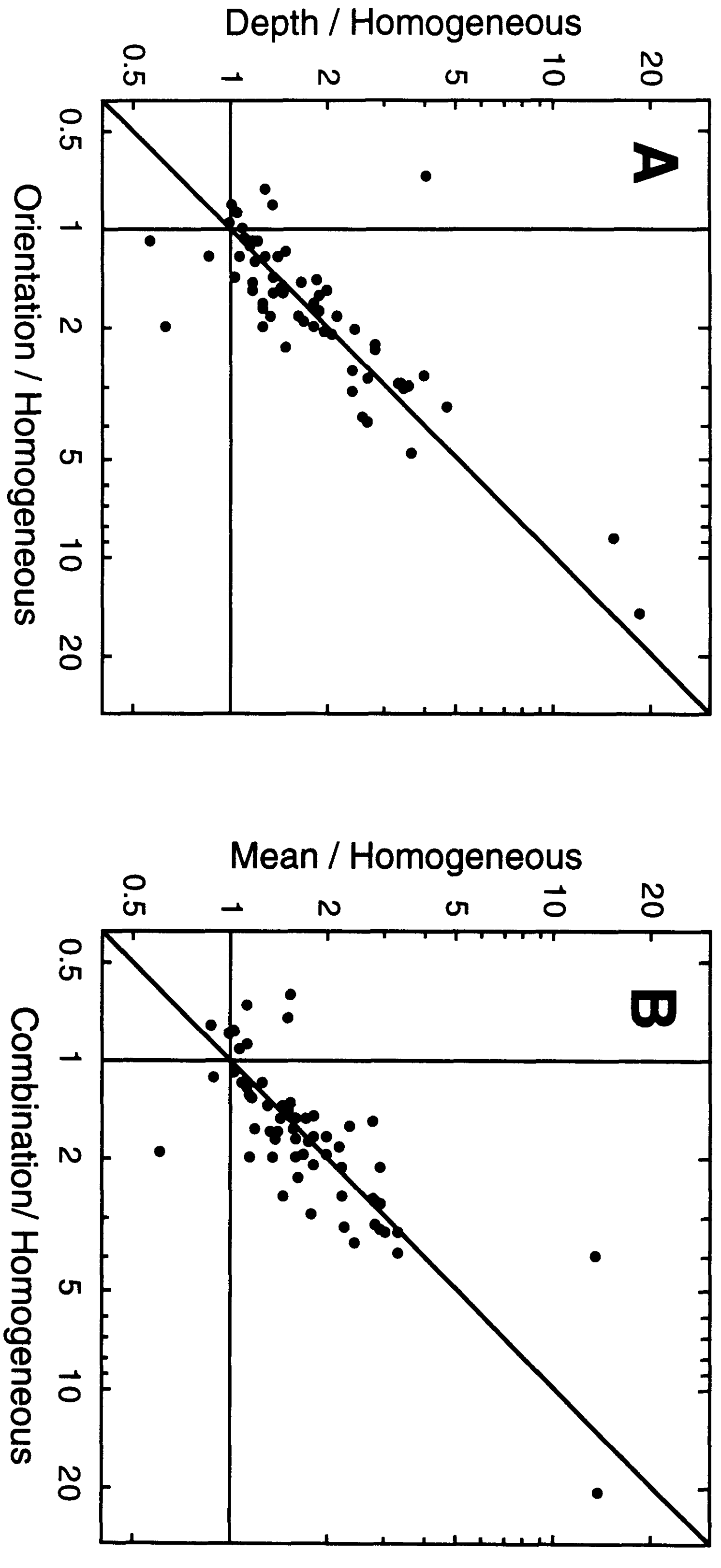
A
Depth / Homogeneous
Orientation / Homogeneous
B
Mean / Homogeneous
Combination/ Homogeneous
0.5
1
2
5
10
20

Figure 12

	depth	**comb.**	**lum.**	**color**
ori.	0.74	0.74	0.51	0.65
color	0.69	0.72	0.38	
lum.	0.59	0.46		
comb.	0.70			

Table 1

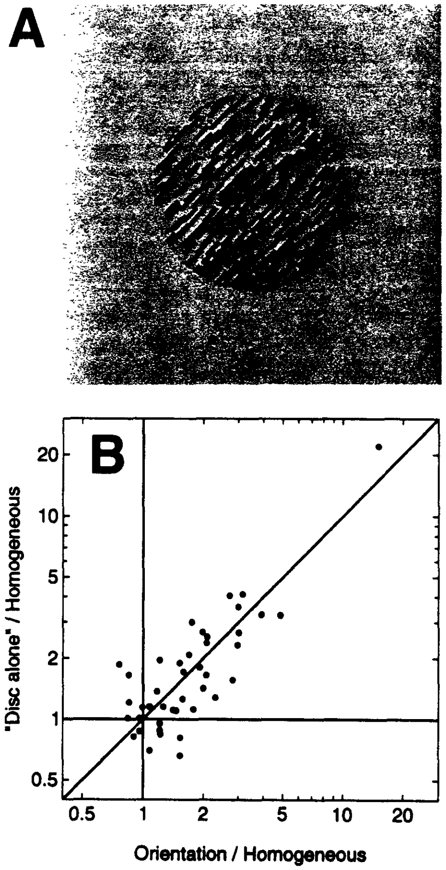

A
B
"Disc alone" / Homogeneous
20
10
5
2
1
0.5
0.5
1
2
5
10
20
Orientation / Homogeneous

Figure 13

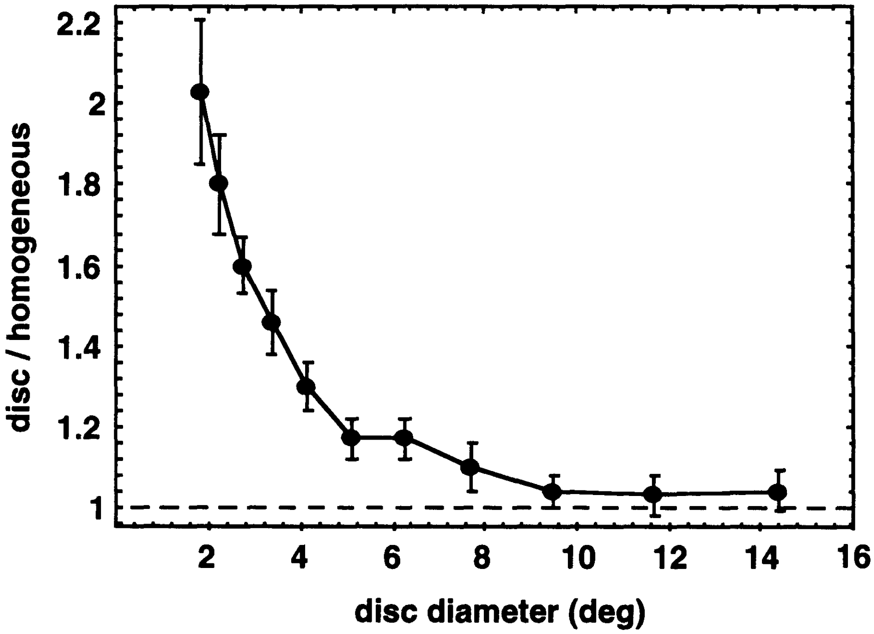

2.2
2
1.8
1.6
1.4
1.2
1
disc / homogeneous
2
4
6
8
10
12
14
16
disc diameter (deg)

Figure 14

Chapter 2

Extra-Receptive Field Modulation: Specificity for Image Structure

Introduction

Given the images impinging on the retinae, the visual system must model the three-dimensional structures of the distal world. When we view scenes such as those depicted in Fig. 1, we gain a strong sense of the physical structure of the surfaces and objects that we infer to have given rise to these images. Achieving this entails far more than image *filtering*, which is merely the process of extracting a subset of the information already present in the input stimulus. Distal structure cannot be found through filtering, because the structures of the distal world modeled so richly in our perception do not in fact exist in the retinal images (Marr, 1982; Kanisza, 1979.) Rather, distal structure must be *inferred* from the traces of contour and texture reflected in the retinal images (Nakayama and Shimojo, 1992). Moreover, because we have a relatively fixed vantage point of a scene at any given moment, the visual system must also make inferences about forms not directly visible, such as the manner in which surfaces complete behind occluding structures (e.g., the occlusion of the wood surface by the tree branches in Fig. 1B; Nakayama et al., 1989; Enns and Rensink, 1994). For the visual system to accomplish these tasks, it must have a sophisticated understanding of the way in which three-dimensional structures in the distal world can lead to the formation of retinal images.

The function of area V1 has long appeared far removed from these concerns. The RFs of V1 neurons were originally described as "line" or "edge" detectors, suggesting that these cells were able to extract from an image only the most elemental units of form (Hubel & Wiesel, 1963). More recent work shows that the V1 RFs are better viewed as spatially restricted filters jointly tuned for orientation and spatial frequency (Movshon et al., 1978; Schiller et al., 1976; Albrecht, 1978). This description, which views the V1 RFs as quasilinear filters capable of extracting the fourier components of local image patches, serves to remove these cells even further from the task of modeling distal world structure. For within this framework, the "lines" or "edges" that may stimulate these cells do so not because they form the contours of surfaces or objects in any perceptual context. Rather, they do so merely because the cells are tuned for the specific two-dimensional spatial frequency content of these stimuli, *regardless of their perceptual context* (K.K. De Valois et al., 1979). The tuning characteristics of V1 RFs for color and binocular disparity are likewise described with simple filters that have no direct connection to perceptual interpretation of distal world structure.

Thus we see that the functional value of the RFs of V1 neurons arises from their specificity in *purifying* particular types of information that exist in images, such as parts of local spatial frequency spectra. The completely separate problem of *synthesizing* from this information a perceptual model of distal world structure (as, for example, in reconstructing the form of the surfaces in Fig. 1) has traditionally been assumed to occur at later stages of visual processing. This view has appeared sensible both because the filter description of V1 seems conceptually complete, and because there was little compelling evidence that V1 neurons could be doing anything qualitatively different from simple image filtering.

The extra-RF modulation recently described by Lamme (1995) and in the previous paper (Zipser et al., 1995) stands in strong contrast to the filter properties of the V1 RFs themselves, however. Unlike the local-patch analysis of the V1 RF, extra-RF modulation has tremendous spatial scope. And in contrast to the RF's role in purifying information about specific visual cues, extra-RF modulation seems to pool information from diverse cues. Although these observations in no way challenge the functional role of the RF itself, they certainly challenge the view that the role of the V1 neuron is *solely* to passively filter local regions of images in a cue-specific manner. Indeed, the massive convergence of signals onto individual V1 neurons that we described could lend these cells greatly increased capacity for visual processing beyond the simple image filtering accomplished by their RFs alone.

In this paper we ask whether extra-RF modulation in area V1 plays a role in modeling distal world structure, such as that of the surfaces we perceive by viewing Fig. 1. Our initial inspiration for this question came from Lamme's suggestion (1995) that extra-RF modulation allows V1 neurons to signal information about the *perceptual context* of stimuli that fall within their RFs. Lamme observed that the extra-RF modulation evoked by a textured figure was fairly constant when the RF of a given cell was at different positions within the figure. However, extra-RF modulation fell off completely when the RF was positioned immediately outside the figure (i.e., when the RF was positioned on the texture background.) For this Lamme proposed that extra-RF modulation could signal the "figure/ground" relationships in a scene. Here we will borrow the idea that extra-RF modulation conveys signals about the larger perceptual context of the small piece of the scene falling within a neuron's RF. We will avoid considering the somewhat abstract concepts of "figure" and "ground," however. Instead, we will focus on the concept of "surface" which is more closely tied to physical form.

Materials and Methods

The basic computer, behavioral, and neurophysiologcial techniques described in the previous paper are also used here. The configuration of textured visual stimuli constitute the main difference between experiments discussed here and those in the previous paper.

Results

As in the previous paper, we here study extra-RF modulation by comparing the responses of V1 neurons to homogeneous texture displays with their responses to displays containing an identical texture pattern over the RF but different patterns outside the RF area. In this paper our purpose lies in studying whether extra-RF modulation is receptive to the *perceived distal structure* of our texture displays. Thus our emphasis here is on varying the perceived distal structure of the display region containing the RF, while at the same time keeping RF texture stimulation the same from trial to trial. A key display that allows us to do this is illustrated in Fig. 2. The display appears as a homogeneously textured field, with the modification that we can manipulate the perceived depth of a band of texture surrounding the RF (i.e., the band of texture between the white dashed lines in Fig. 2; dashed lines are not in the actual display). The texture beneath the RF and in the surrounding background will typically remain at zero disparity.

In the case where we cause the texture band surrounding the RF to have the same binocular disparity as the other regions of the display, we simply generate a standard homogeneous texture display. In Fig. 3A we illustrate this display from an aerial perspective (top), and schematically from a side view (middle). At the bottom of Fig. 3A we illustrate the average response profile of one V1 neuron (cell *z*) to stimulation with this display in one-hundred and fifty trials. Cell *z*'s response to this display showed an initial vigorous burst of activity in response to texture onset, followed by a much diminished response rate for the remainder of the 250 ms texture display interval.

Moat: We could alter the perceived distal structure of the display by simply causing the band of texture surrounding the RF to appear directly back in depth from the remaining texture (through 0.14° uncrossed horizontal disparity.) We refer to this receded region as a "moat," illustrated in the top and middle of Fig. 3B. As seen from these schematic illustrations, with establishment of the "moat" the RF no longer appears positioned on a large textured field, but rather appears to be positioned on a small square surface isolated from the textured background by the "moat." In the experiment, moat depth was only apparent through binocular disparity cues, although we provide some shading cues to depth in Fig. 3 for schematic purposes.

We stimulated cell *z* with the moat display in thirty trials (randomly interleaved with those for the homogeneous texture display.) At the bottom of Fig. 3B we illustrate the cell's resulting average response profile. The initial response of the cell for the moat display was nearly identical to the response for the homogeneous texture display. However, approximately 100 milliseconds after texture onset, the response rates diverged, with the moat display causing the cell to maintain a more vigorous response rate than did the homogeneous texture display (gray shading of the response profile.) Thus, we see that the moat display evoked extra-RF modulation of the same nature as we have seen with the various disc displays in the previous paper.

Frame: We could also modify the display in Fig. 2 in a different way by having the texture band surround the RF appear nearer in depth than the remaining texture (through 0.14° crossed horizontal disparity.) In this case the perceived distal structure (Fig. 3C top and middle) is completely different from the moat display. In the frame display the RF appears positioned not on a small textured surface, by on a large textured surface *continuous* with the textured background, as though a narrow textured "frame" were merely floating above and partially occluding the homogeneous texture display in Fig. 3A. In the bottom of Fig. 3C we illustrate cell *z*'s average response rate to the frame display from stimulation in thirty randomly interleaved trials. The results stand in strong contrast to the response to the moat display, because cell *z*'s response to the frame display very closely follows that to the homogeneous texture display. [The only exception to this is the "off" response after extinguishing of the texture display. The response of cells to the offset of texture displays falls outside our 100 to 300 ms analysis interval and is not of concern to us here.]

In summary, in Fig. 3 we examined the response of a V1 neuron to three different texture displays, two of which (the homogeneous display and the frame display) have common perceived distal structure in that the RF appeared positioned on a large texture surface. For both of these displays the cell showed the same muted response. In contrast, for the moat display, in which the RF appears positioned on a small texture surface, the cell responded vigorously.

Population analysis of the moat/frame asymmetry for evoking extra-RF modulation

Remarkably, this asymmetry of effect for the moat display compared to the frame display was highly consistent across the seventy single and multi-unit sites that we studied with these stimuli. We demonstrate this in Fig. 4 , which illustrates histograms of extra-RF modulation ratios for these recording sites. In Panel A we show the ratio values for moat response / homogeneous response. Extra-RF modulation ratio values in this case fall consistently above 1.0, indicating that neural responses for the moat display generally exceeded those for the homogeneous texture display. In Panel B we show the ratio values for frame response / homogeneous response. In contrast to the moat case, here the extra-RF modulation ratio values cluster near 1.0, indicating that for the frame display, neurons responded in a similar fashion as to the homogeneous texture display.

Perturbations in moat and frame displays that retained the essential character of their perceived distal structure evoked qualitatively similar results to those just described. For example, the asymmetry in effect of moat and frame displays for evoking extra-RF modulation did not depend on having the displays centered at zero disparity (the standard case), but was equally evident when we moved all texture displays back in depth relative to the fixation spot (i.e., by manipulating binocular horizontal disparity cues; data not shown.) Furthermore, we could vary the magnitude of the "moat" and "frame" disparities to larger or smaller values than our +/-0.14° standard without upsetting the basic effect. In summary, the asymmetry in effectiveness of the moat versus the frame displays for evoking extra-RF modulation is a consistent and robust result.

Moats and Frames delineated by two cues

Given the pronounced difference in extra-RF modulation elicited by moat and frame displays, it is interesting to ask what would happen if we added a second cue for differentiating these displays from the homogeneous texture display. We illustrate a display that allows us to do this in Fig. 5. Here the texture band that surrounds the RF has texture orientation orthogonal to the remaining texture in the display. As before, we will vary the apparent depth of this texture band in order to manipulate the perceived distal structure of this display.

First we describe the effects when all texture regions of the display in Fig. 5 are of the same binocular disparity. In this way we arrive at the *orientation-band* display schematized in the top and middle of Fig. 6A. We may point out that the perceived distal structure of this display is somewhat ambiguous. To most observers, the display appears bi-stable, taking on the appearance of either a moat or a frame display from moment to moment. We illustrate the average response profile of cell z to this display at the bottom of Fig. 6A. [This and the remaining two response profiles in Fig. 6 are based on thirty trials each, all randomly interleaved with trials presented in Fig. 3.] We see that stimulation with the orientation-band display caused the cell to maintain a firing rate elevated above that for the homogeneous texture display (gray shading of response profile.) The difference in response between homogeneous and orientation-band displays was only about half as large as between the moat display and the homogeneous display seen in Fig. 3; nevertheless, the difference was highly significant ($p < 0.01$).

Given that for this cell the orientation-band alone evokes extra-RF modulation, we can now ask what occurs when we add the same orientation cues to the moat and frame displays to generate *moat & orientation* and *frame & orientation* displays (schematized in the top and middle of Figs. 6B and 6C respectively.) We may point out that as far as the perceived distal structure of these displays is concerned, the addition of the orientation cues in no way alters the essential character of the "moat" or "frame" structures defined by binocular disparity cues. The results of stimulation with these displays for cell *z* (illustrated in the bottom of Figs. 6B and 6C) show that addition of the orientation cue has no effect on the response to moat or frame displays. In neither case do the response rates during our standard 100 to 300 ms analysis interval differ significantly from the responses to the normal moat and frame displays depicted in Fig. 3.

Across the subset of forty-four recording sites that we studied with both the displays in Fig. 3 and in Fig. 6, we found the same trend of results as with cell *z*. The orientation-band display evoked a modest level of extra-RF modulation (the median extra-RF modulation ratio equals 1.12.) Yet the responses evoked by the moat & orientation display were in the median case almost exactly the same as those evoked by the normal moat display (median ratio of the former to the latter equals 1.01). And finally, the frame & orientation display actually evoked slightly *reduced* responses compared to the normal frame display (median ratio of the former to the latter equals 0.92). Thus in summary, adding the orientation cue to moat and frame displays had no enhancing effect of the neural response to these displays, despite the fact that the orientation-band itself could evoke at least limited extra-RF modulation (as exemplified by cell *z*.)

Other disparity displays

One characteristic difference between moat and frame displays is that in the former the texture on which the RF rests is surrounded by texture of relatively farther depth, whereas in the latter the surrounding texture is relatively nearer in depth. Is this lone characteristic enough to account for the different extra-RF modulation results obtained with moat and frame displays?

In this section we address this question by comparing the extra-RF modulation evoked by the *disparity disc* display (described in the previous paper) and the *disparity-hole* display. These displays are schematized at the top of Figs. 7A and 7B respectively. In the disparity disc display the RF appears positioned on a disc floating above a textured background. In the disparity-hole display, the RF appears positioned on surface seen through a disc-shaped hole. Only binocular disparity cues reveal these depth relationships in the actual stimuli. We illustrate the average response profiles of cell *z* to these displays in the lower part of Fig. 7. In each case we compare the responses to activity evoked by a homogeneous texture display of appropriate binocular disparity to match RF texture. In both cases we see that cell *z*'s response exceeds that to the appropriate homogeneous texture display (gray shading of response profiles). Thus cell *z*, which showed a clear asymmetry in effectiveness of moat versus frame displays for evoking extra-RF modulation, nonetheless showed clear extra-RF modulation for both disparity disc and disparity-hole displays.

We found this result to be consistent across the seventy individual V1 neurons that we studied with these displays (including sixty-five already discussed in the previous paper). In Fig. 8 we demonstrate this point with extra-RF modulation ratio histograms for both disparity disc and disparity-hole displays. These histograms show that the strong tendency of these two displays to evoke extra-RF modulation is indistinguishable across the cells we studied.

In summary, we cannot account for the difference in effect of moat and frame displays simply by the relative disparity of surrounding texture. Rather, the contrast in results between these displays and the disparity disc and disparity-hole displays shows that we must take into account the broader structure of the extra-RF pattern.

The ring display

In the previous section we found that disparity contrast, without regard to overall spatial pattern, cannot explain the differential effects of moat and frame displays. In this section we ask, is it possible to mimic the results of the frame display with a stimulus that contains no explicit binocular disparity cues?

We illustrate the display that we will use for this purpose, the *ring display*, in Fig. 9. The ring display consists of a homogeneously textured field (the same as the homogeneous texture display) on which is drawn a black circular contour of the same binocular disparity and luminance as the black texture bars. The RF of a given V1 neuron under study is positioned in the center of the "ring." The perceived distal structure of this display resembles that of the frame display, in that texture within the ring contour appears to form a *continuous* surface with the surrounding texture (as though the ring were merely resting on top of the texture surface.)

In Fig. 10 we investigate for cell *z* the extra-RF modulation evoked by the ring display and normal disc displays that we described in the previous paper (ring inner diameter and disc diameter was 3.6°). The top response profile in this Figure shows the response to the homogeneous texture display, and the second response profile shows the response to the ring display; the two response profiles are virtually indistinguishable. Results for the textured discs defined by orientation, disparity, color, luminance, or a combination of these cues are illustrated in the next five response profiles. In contrast to the ring display, each of these displays caused the cell to retain a vigorous response rate relative to the homogeneous texture display (gray shading of response profiles.) Thus, for cell *z* the ring display acted just like the frame display.

The effect of the ring display was also qualitatively different from the disc displays across the population of forty-four V1 neurons that we tested with these displays. We make this point in Fig. 11. Here we illustrate, in Panel A, a histogram of the extra-RF modulation ratios for the ring display (i.e., ring response / homogeneous display response) and, in Panel B, the extra-RF modulation ratios for the disc displays. The median extra-RF modulation ratio for the ring display is 0.97, whereas across the five disc displays it is 1.48. Thus, while the disc displays consistently evoked greater response rates than the homogeneous display, the ring display acted like the frame display in evoking response rates similar (in the median) to the homogeneous texture display.

Discussion

In order for us to achieve our rich perception of three-dimensional structure of the distal world, the visual system must somehow model this structure. For the visual system to accomplish its tasks, it must have a sophisticated understanding of the way in which three-dimensional structures in the distal world can lead to the formation of retinal images. There must of course be a neural correlate of this process. From this it follows that in some way, individual neurons must change their firing depending on the characteristics of perceived distal structure. In this paper we have asked whether neurons in area V1 are involved in this process in a way that we may observe.

Although the experiments that we describe in this paper were designed for the purpose of studying whether extra-RF modulation in area V1 plays a role in modeling distal world structure, the data that we collected in the course of these investigations are of interest in their own right. We will thus first summarize all the results without concerning ourselves with larger questions of perception. Once we have reviewed the data, we will return to a discussion of how they may reflect the process of modeling distal world structure.

Summary of results

Disc displays

In the previous paper we used disc displays to show that a diverse set of visual cues may evoke extra-RF modulation. A striking result was that some neurons displayed extra-RF modulation that was invariant over the various cues or their combination. Most cells were receptive to a variety of disc-defining cues. Yet despite the lack of clear specificity for visual cues, extra-RF modulation was spatially tuned for disc diameter.

Moat display

Because our initial experiments always used displays in which a textured disc segmented from a large background of heterogeneous texture, it was certainly conceivable that extra-RF modulation in fact *required* a large heterogeneous background of this type in order to be activated. When we first began experimenting with the moat display it was thus an open question as to whether or not this stimulus, which features only a narrow band of disparity contrast around the RF, could in fact evoke extra-RF modulation. The results in Figs. 3B and 4 showed that the moat display does indeed elicit consistently more vigorous response rates than the homogeneous display. This thus leads us to the important finding that although extra-RF modulation has large spatial scope (as we demonstrated in the previous paper), its modulatory effect may be elicited by fairly small modifications of the stimulus display.

Frame displays

Up to this point we have examined texture displays that evoke stronger response rates from V1 neurons than does the homogeneous texture display. In contrast, our experiments with the frame display yielded the important finding of a stimulus configuration that did *not* evoke extra-RF modulation in the manner that we have previously seen. Rather, neural responses to the frame display were approximately the same (or even slightly less) than the responses to the homogeneous texture display. We found this result to be consistent and robust over a number of perturbations to the frame display (Figs. 3C, 4, and 5). This result is potentially very exciting because it suggests that extra-RF modulation maintains *specificity* for stimulus configuration of a type not previously described.

A simple explanation for the failure of the frame display to evoke extra-RF modulation in the same manner as does the moat display is that extra-RF modulation is simply *insensitive* to the disparity difference between the "frame" and the background texture. We are able to rule out this possibility by comparing the results for the frame & orientation display with the results for the orientation-band display. While the orientation-band display could evoke a modest increase in response rate over the homogeneous texture display (Fig. 6A), this extra-RF modulation *disappeared* for the frame & orientation display (Fig. 6C). Clearly, if extra-RF modulation had merely been insensitive to the difference in disparity between "frame" and background, the neural response for the frame & orientation display should have been the same as to the orientation-band display. Our results therefore indicate that the "frame" disparity was indeed registered by the mechanisms underlying extra-RF modulation, and in some way canceled the effect of the orientation cue. Extra-RF modulation's specificity for stimulus configuration may thus be of a rather sophisticated nature.

Disparity-hole and ring

Although the frame display elicits different results from the moat and disc displays, it is not clear from what we have discussed so far exactly what makes the frame display act differently. There is no a priori reason that the pattern heterogeneity of the "frame" should be any less potent than that of the "moat." Could the *near* quality of the "frame" surrounding the RF cause extra-RF modulation to treat the frame display differently? Or, is the overall narrow structure of the "frame" important as well?

The disparity-hole display allowed us to address this question. The disparity-hole display, like the frame display, features relatively *near* texture around the RF. Yet with this display (Fig. 7B) we found normal extra-RF modulation, quite unlike the result with the frame display. Thus, something about the overall structure of the frame display is important for the results that we recorded, further supporting the notion that extra-RF modulation's specificity for stimulus configuration is of a sophisticated nature.

The ring display is of interest because it demonstrates that the mere presence of a contour around the RF is not sufficient to evoke extra-RF modulation.

Lamme's result

Lamme (1995) found relatively constant extra-RF modulation when a cell's RF lay within a textured figure (a 4° x 4° square). However, when the RF was positioned directly outside the figure, extra-RF modulation disappeared. The result is curious, as we might expect that the textured figure immediately adjacent to the RF could activate extra-RF modulation. The result favors the view that extra-RF modulation involves more than a simple lateral inhibition mechanism.

In summary, through the experiments in this two part series and previously by Lamme, we have seen extra-RF modulation to have an interesting collection of characteristics. We have seen that it is a phenomenon of broad spatial scope which is nonetheless sensitive to fairly small scale perturbations of the stimulus display. It may be evoked by a wide variety of visual cues, and under certain stimulus conditions can respond invariantly over individual cues or their combination. Yet under other stimulus condition, perfectly salient visual cues will fail to evoked extra-RF modulation--and may even block its effects. The complexity and apparent flexibility of this phenomenon appears quite different from the simple filter properties of the V1 RFs themselves. We now turn to the task of finding a functional explanation for these results.

Interpreting V1 extra-RF modulation

Hypothesis for extra-RF modulation's functional role

For the purpose of discussion, we will attempt to explain our results with the following conjecture: extra-RF modulation reflects an analysis of the stimulus display in terms of the *three-dimensional distal structure* inferred to have given rise to the display. For displays like ours, in which the RF appears positioned on a textured *surface*, we postulate that extra-RF modulation is triggered by apparent *discontinuities* in the surface itself. As a working hypothesis, we propose that the strength of extra-RF modulation depends on how closely positioned the discontinuities are to the RF.

We can attempt to use these conjectures to explain our results as follows.

Consider a disc shaped surface with a V1 neuron's RF in the center. As disc diameter increases, the discontinuities in the surface created by the disc contour move farther from the RF, thereby reducing the strength of extra-RF modulation. In this way we explain the spatial tuning of extra-RF modulation for disc size describe in the previous paper.

Next, consider that we choose a disc surface of a fixed size and position it over different backgrounds. Depending on the background, the disc surface may be visible by a wide variety of cues; but as long as the disc can be easily segmented as a surface, extra-RF modulation based on the surface form of the disc should remain the same for the given disc diameter. In this way we explain the cue-invariant properties of extra-RF modulation that we observed.

Next, consider the moat display. The RF appears positioned on a small surface, surrounded by the "moat" discontinuity. From the perspective of surface discontinuity, it is irrelevant that the "moat" is narrow and does not fill up the entire background. The important thing is that the "moat" creates discontinuity in the RF's surface, leading to extra-RF modulation. In this way we explain how spatially restricted manipulations of the stimulus display can evoke modulation as can disc displays. The same holds true for the moat & orientation display.

Next, consider the frame display. From the perspective of three-dimensional distal structure, the RF appears positioned on a homogeneous texture surface that merely happens to be partially occluded by the "frame." As the "frame" does not appear to create any discontinuities in the surface that harbors the RF, extra-RF modulation is not evoked. The same holds true for the frame & orientation display.

Next, consider the disparity-hole display. In this case, the extent of the surface on which the RF is positioned is obscured because of the limited size of the hole. The RF's surface may thus be judged to be just large enough the fill the bottom of the hole, or at least to be moderate in size compared to the homogeneous texture display. In this way, inferred discontinuities in the RF's surface will evoke the extra-RF modulation, in contrast to the result with the frame display.

Next, consider the ring display. As the ring can be seen to merely rest upon the homogeneous texture display, it does not thus appear to cause any discontinuity in this surface. For this reason, it fails to evoke extra-RF modulation, just as we found with the frame display.

Finally, consider Lamme's result. In Lamme's experiment, the textured figure appeared to occlude its continuous textured background. Thus, when the RF was on the textured background (even immediately *outside* the texture figure), no extra-RF modulation was evoked because the background surface had no discontinuities of its own.

In summary, we find the hypothesis that extra-RF modulation serves to detect discontinuities in the inferred structure of distal surfaces to be useful for explaining the data within a common framework. The hypothesis is exciting because it implies that a rather advanced analysis of form is reflected in the neural activity of area V1. There are of course a number of caveats that we must now discuss.

Discussion of the surface hypothesis

An obvious concern raised by the surface hypothesis is that not all visual stimuli are conveniently described with a surface description. For example, clouds, rain, and hair do not have the same sense of surface as our textured displays. We thus might question whether extra-RF modulation would play a role specifically tied to surface form. One possible solution to this problem is that extra-RF modulation's function (and its interpretation by other neural centers) may be contingent on the nature of the distal structure within the RF of the recipient neuron. Thus, if the RF of a cell is positioned on a surface, extra-RF modulation will deliver signals relevant to this condition. In other cases, it may operate on a different set of rules--or not at all.

A related concern raised by the surface hypothesis is that it does not provide a complete description of how surfaces are represented. There are many aspects of representing surfaces that we have not discussed (such as transparency, curvature, and rigidity) that are salient perceptually. It is not presently clear how these parameters could be represented in the extra-RF modulation framework. But once again, flexibility in how extra-RF modulation is computed and interpreted could substantially increase its functionality in representing these different surface forms.

Alternative explanation

The surface hypothesis discussed above is perhaps one of the most interesting interpretations of the data, but of course we must be open to the possibility that extra-RF modulation plays a far more limited role in visual analysis. While with the surface hypothesis we proposed that the properties of extra-RF modulation that we observed reflects a generalized analysis of surfaces, it is certainly conceivable that extra-RF modulation in fact consists of an ensemble of relatively simple mechanisms that yield the results we found but have no larger power. Future experiments will be needed to distinguish between these possibilities.

Figure Legends

1 Examples of surfaces. **A** shows textured surfaces forming an object. **B** shows a textured surface occluded by a tree.

2 Configuration of a texture display in which a band of texture surrounding the RF may vary in apparent depth through binocular disparity cues.

3 Homogeneous, moat, and frame displays. **A** shows schematic illustrations of the *homogeneous* texture display (top and middle) and the response of one neuron (cell *z*) to this display (bottom). **B** shows data for the *moat* display in the same format. **C** shows data for the *frame* display in the same format.

4 Extra-RF modulation ratios for frame and moat displays. Upper panel shows extra-RF modulation ratios for moat display. Lower panel shows extra-RF modulation ratios for frame display.

5 Configuration of a texture display in which a band of texture surrounding the RF may vary in apparent depth through binocular disparity cues. In this case the band of texture has orthogonal orientation to the remaining texture.

6 Orientation-band, moat & orientation, and frame & orientation displays. **A** shows schematic illustrations of the *orientation-band* texture display (top and middle) and the response of one neuron (cell *z*) to this display (bottom). **B** shows data for the *moat & orientation* display in the same format. **C** shows data for the *frame & orientation* display in the same format.

7 Comparison of floating disc and disparity-hole displays for cell *z*. In each case, the cell displays extra-RF modulation.

8 Extra-RF modulation ratios for floating disc and disparity-hole displays. Upper panel shows extra-RF modulation ratios for floating

disc display. Lower panel shows extra-RF modulation ratios for disparity-hole display.

9 Configuration of ring display.

10 Response of cell z to homogeneous texture display, ring display, and the five disc displays described in the previous paper.

11 Extra-RF modulation ratios for ring display and disc displays. Upper panel shows extra-RF modulation ratios for ring display. Lower panel shows extra-RF modulation ratios for disc displays.

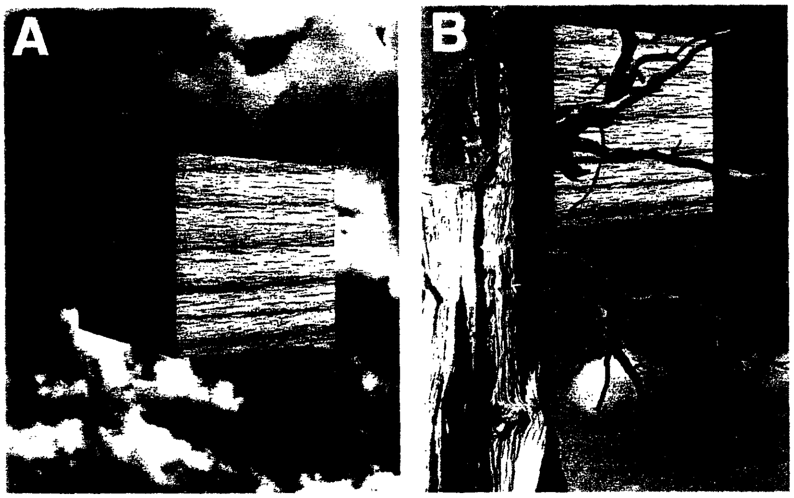


Figure 1

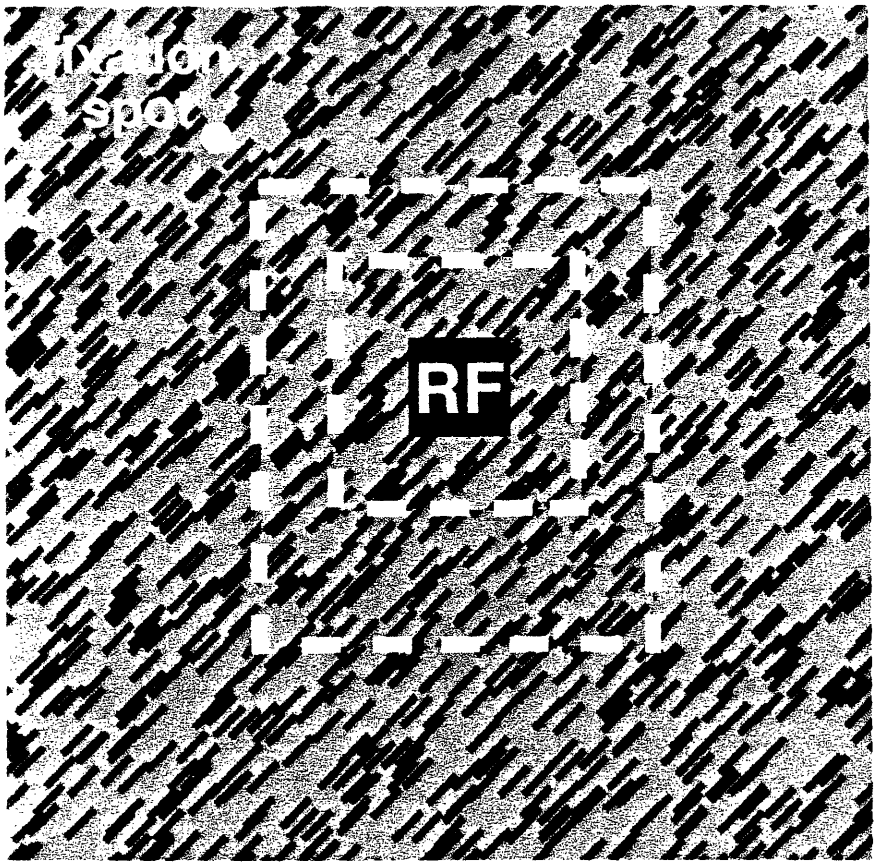


Figure 2

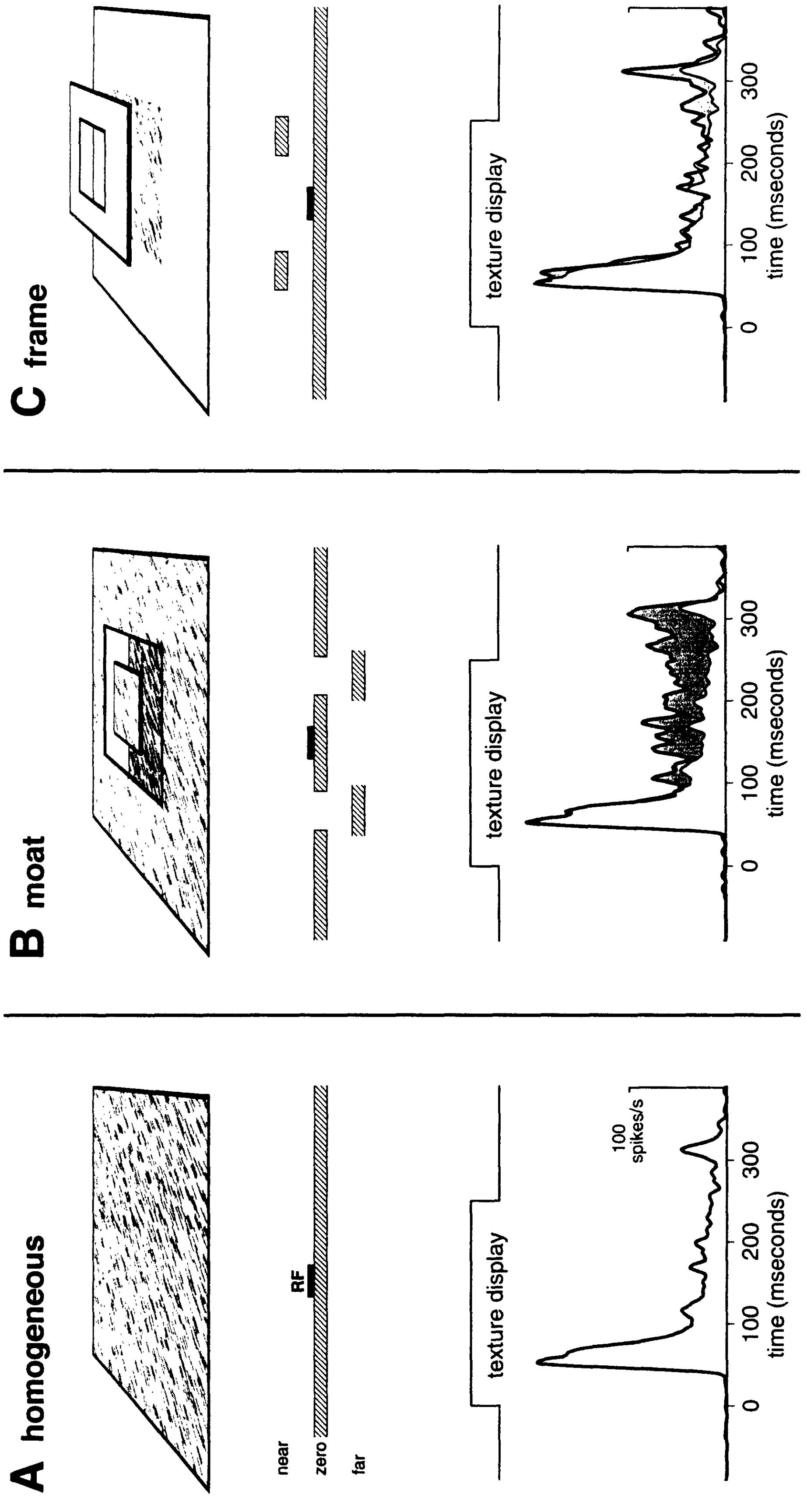

A homogeneous
near
zero
far
RF
texture display
100 spikes/s
0
100
200
300
time (mseconds)
B moat
texture display
0
100
200
300
time (mseconds)
C frame
texture display
0
100
200
300
time (mseconds)

Figure 3

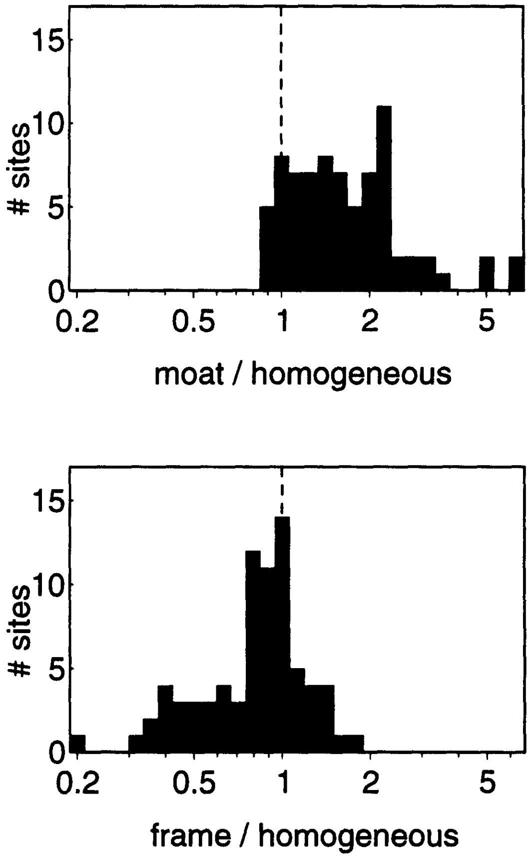

15
10
5
0
sites
0.2
0.5
1
2
5
moat / homogeneous
15
10
5
0
sites
0.2
0.5
1
2
5
frame / homogeneous

Figure 4

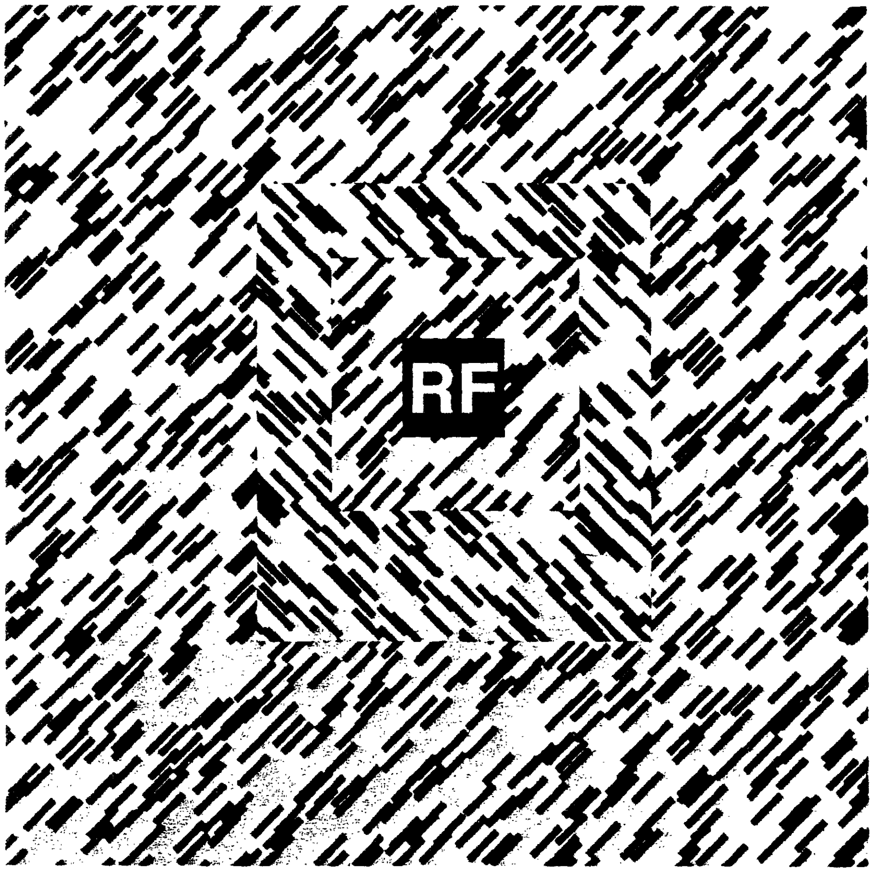


Figure 5

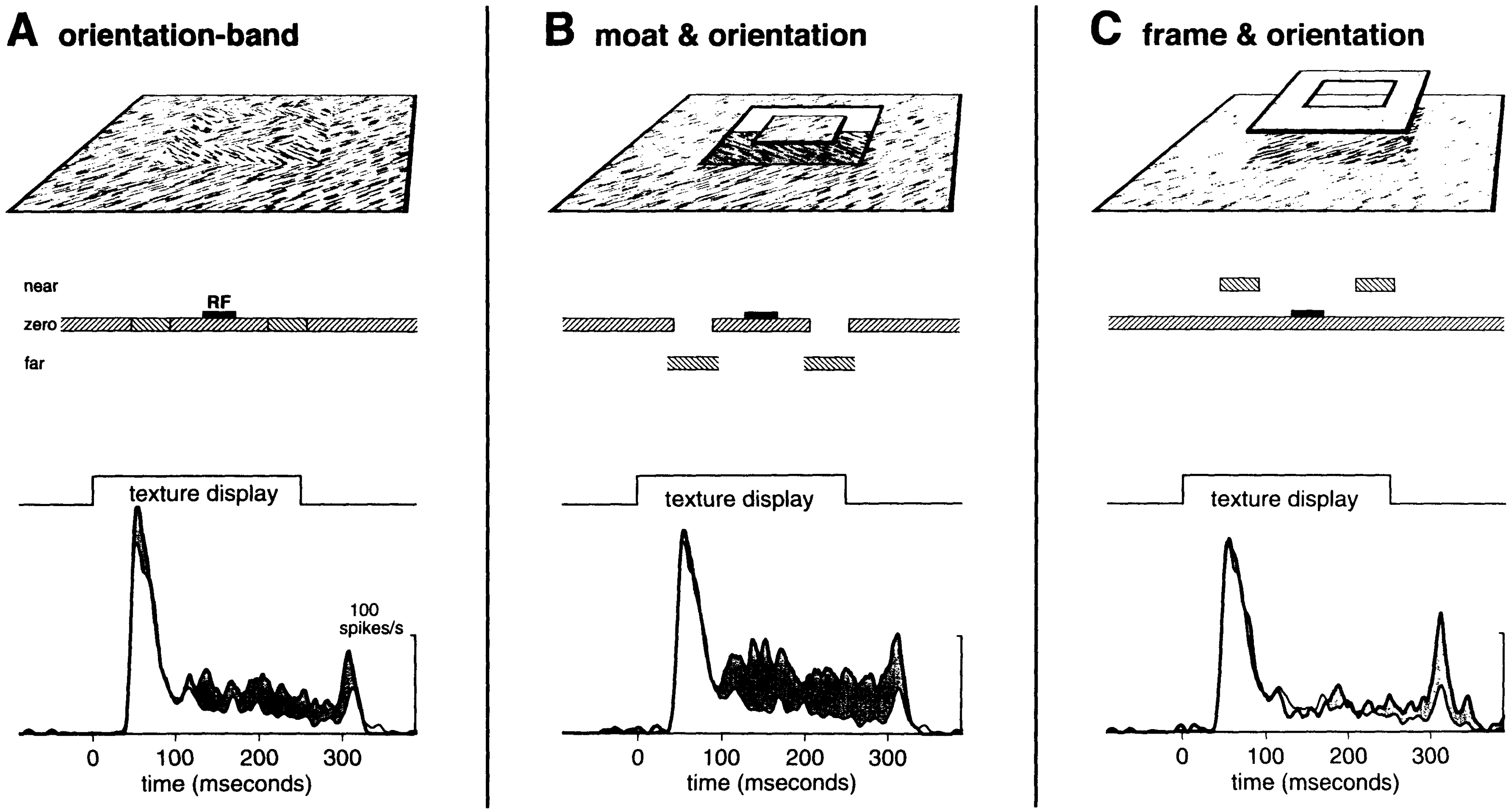

A orientation-band
B moat & orientation
C frame & orientation
near
RF
zero
far
texture display
texture display
texture display
100
spikes/s
0
100
200
300
time (mseconds)
0
100
200
300
time (mseconds)
0
100
200
300
time (mseconds)

Figure 6

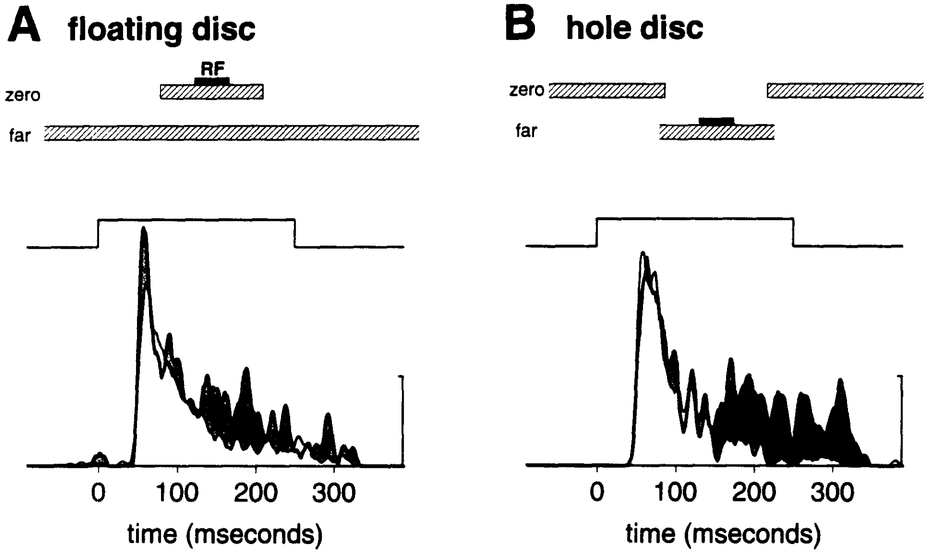

A floating disc
RF
zero
far
0
100
200
300
time (mseconds)
B hole disc
zero
far
0
100
200
300
time (mseconds)

Figure 7

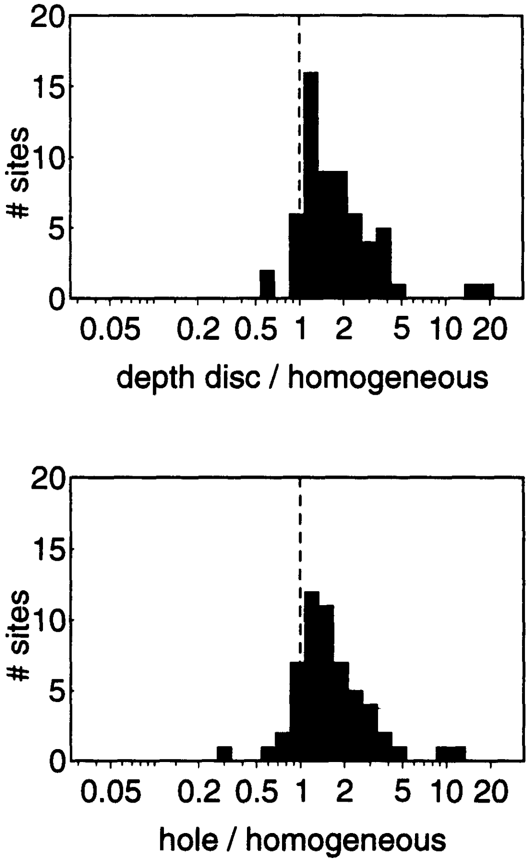

20
15
10
5
0
sites
0.05
0.2
0.5
1
2
5
10
20
depth disc / homogeneous
20
15
10
5
0
sites
0.05
0.2
0.5
1
2
5
10
20
hole / homogeneous

Figure 8

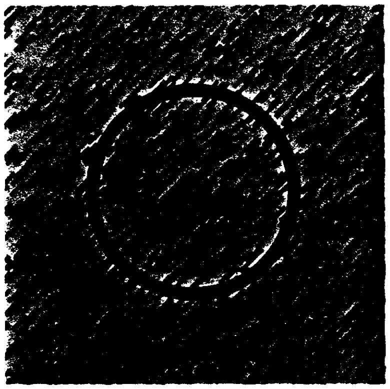


Figure 9

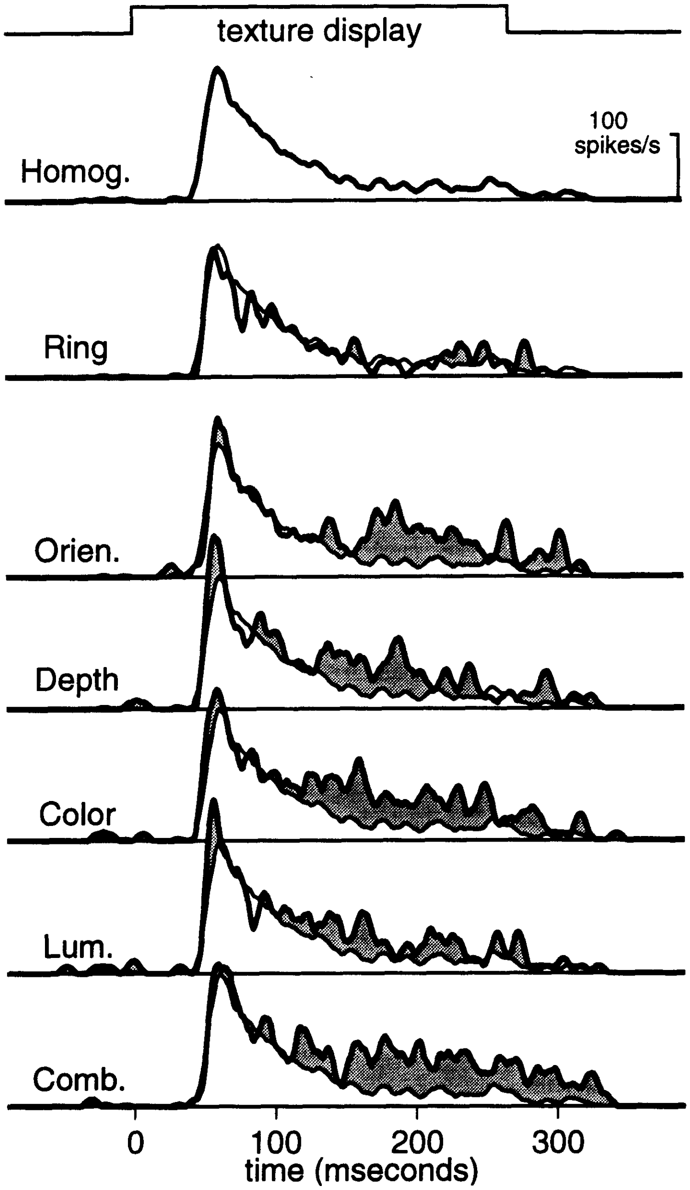

texture display
100 spikes/s
Homog.
Ring
Orien.
Depth
Color
Lum.
Comb.
0
100
200
300
time (mseconds)

Figure 10

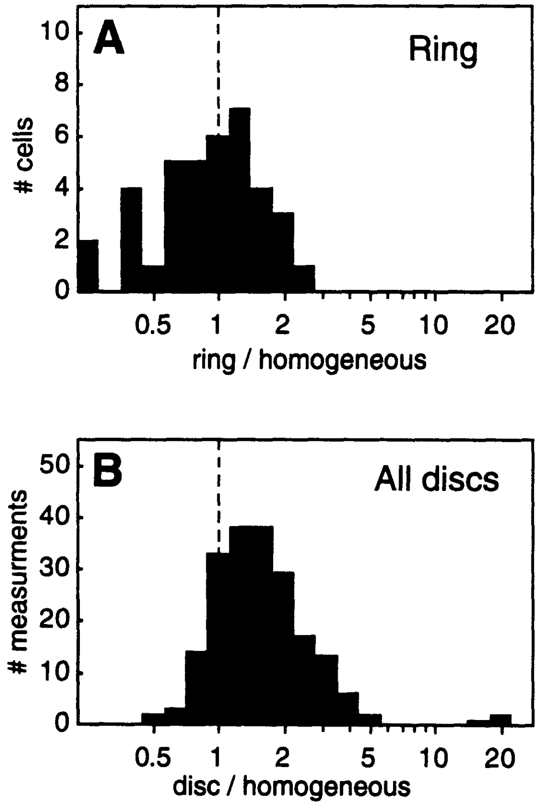

A
Ring
10
8
6
4
2
0
cells
0.5
1
2
5
10
20
ring / homogeneous
B
All discs
50
40
30
20
10
0
measurments
0.5
1
2
5
10
20
disc / homogeneous

Figure 11

Chapter 3

Extra-Receptive Field Modulation: Temporal Properties

Introduction

An interesting feature of the extra-RF modulation that we have described in the previous two papers of this series is the temporal delay in the onset of this effect. Our standard way of studying extra-RF modulation has been to flash a textured display onto the display as a monkey fixates. The typical result that we have observed is that V1 neurons respond to texture onset with a burst of activity. Given identical stimulation of the RF, it is generally only after this initial burst of activity (i.e., 30 to 50 milliseconds after the initial neural response) that the effects of extra-RF modulation become evident. We have seen that for the homogeneous textured display, neural response rates often fall off after the initial burst of activity, whereas with other displays (such as textured discs) the neurons retain more vigorous response rates.

The delay in the onset of extra-RF modulation is potentially interesting for a number of reasons. One reason is because it suggests that extra-RF modulation is truly a distinct neural process from the normal RF functioning of a V1 neuron. For in contrast to the delay in expressing extra-RF modulation, V1 neurons display their tuning specificity for visual stimuli with their first action potential responses to visual stimulation (Celebrini et al, 1993). A second reason why the delay in the onset of extra-RF modulation is interesting is because it raises the possibility that time-consuming and complex computations must occur before this phenomenon is expressed, as compared to the relatively simple filter operations of the V1 RFs themselves. The extra 30 to 50 milliseconds before extra-RF modulation appears is enough time for visual signals to travel up to extra-striate cortical areas and back to area V1. Thus, the temporal delay raises the possibility that extra-RF modulation could reflect the more advanced neural processing of these higher visual areas. A third reason why the delay in the onset of extra-RF modulation is interesting is because it raises a possibility for the visual system to separate signals pertaining just to RF stimulation (found in the initial response phase) from signals that also reflect analysis of the area outside the RF (found in the later response phase.)

However, before we may draw any firm conclusions about extra-RF modulation based on its timing, we must confront a difficulty inherent in our previous study of this phenomenon. This difficulty is that our previous experiments do not allow us to rule out the possibility that the delay of the onset of extra-RF modulation is somehow influenced by initial burst of activity to texture onset. For example, perhaps the initial burst of activity at texture onset somehow artificially *delays* the appearance of extra-RF modulation. Alternatively, extra-RF modulation might *require* the driving force of the initial burst of activity in order to be activated, meaning that it could not occur without the recent appearance of an RF stimulus. In either of these cases, the temporal delay of extra-RF modulation would depend in rather trivial ways on the fact that the RF of the given neuron had just been hit by visual stimulation. In order to show that the delay in the onset of extra-RF modulation is a characteristic feature of the phenomenon itself, and not merely a trivial side effect of RF stimulation, we need to show that this delay is independent of when the RF itself was first stimulated. The goal of this paper is to study the timing of extra-RF modulation independently from the temporal dynamics of the response to direct RF stimulation.

Results

Fig. 1 compares response profiles for stimulation of V1 neurons with moat and homogeneous texture displays. The data shown are an average' of the responses recorded at three different multi-units sites. The initial response to texture stimulation was the same for moat and homogeneous texture displays. However, approximately 100 milliseconds after texture onset, the response rates diverge. The response to the homogeneous display declines precipitously, whereas the response to the moat display remained at a more vigorous level. We highlight the difference in response with the gray shading of the response profile composite.

The extra-RF modulation evident for the response to the moat display (compared to the homogeneous texture display) in Fig. 1 appears approximately 50 milliseconds after the initial response to texture stimulation. In order to determine if this result is in some way tied to the cell's initial burst of activity, we need to try to evoke extra-RF modulation when the RF is receiving a steady input stimulus.

How may we try to accomplish this? One simple way is to use a two-step procedure in which we first present a homogeneous texture display (thereby generating the initial burst of neural activity), and the subsequently modify only the extra-RF stimulus, leaving RF stimulation the same. We can contrast these results to the response when the homogeneous texture display remains unchanged throughout the entire period.

In Fig. 2A we illustrate the response of the same cells as in Fig. 1 to a long duration stimulation with the homogeneous texture display. We see that after the initial burst of activity, the response rate settles to a steady state of activity. In Fig. 2B we illustrate the response to the two step procedure, starting with homogeneous texture display, and then, 150 milliseconds later, manipulating just the background texture to change to the moat display. The response profile is shown in standard composite with the response profile of Fig. 2A. The clear result is that approximately 100 milliseconds after the display changed to the moat configuration, the response rate *rebounds* to a more elevated level of activity (indicated by the gray shading of the response profile).

This result is important because it indicates that extra-RF modulation need not be triggered by an initial burst of activity. Rather, the results show that extra-RF modulation may be triggered even when neurons have achieve a steady state of firing from constant RF stimulation.

In Fig. 3 we look more closely at the timing of extra-RF modulation under the normal one-step texture stimulation and in the two-step case. Figs. 3A and 3B duplicate the two-step and one-step results from Figs. 1 and 2. In Figs. 3C and 3D we illustrate the *difference* in response rates for the homogeneous texture displays and the moat displays for the two- and one- step paradigms respectively. The "difference" response profiles in Figs. 3C and D are fairly similar. Importantly, both show the beginning of activation (i.e., extra-RF modulation) at the same time (approximately 100 milliseconds after the change in texture background.

The two-step texture display procedure gives us the opportunity to perform an interesting control not available with the one-step procedure. It is possible to start with a homogeneously textured display and replace some of the extra-RF texture with additional randomly generated texture of the same type. Replacing texture with texture of the same type of course has no steady-state perceptual consequences--for the new texture merges into the old texture. However, the initial appearance of the new texture produces a potent transient stimulus nonetheless for V1 cells with their RFs over the new texture nonetheless. We can show this by placing the RF of a multi-unit site over the part of the display that we change with the two-step procedure. In Fig. 4A we show the response of the multi-unit site to the two step procedure, starting with the homogeneous texture display (which produces the first transient of activity), and then changing texture in the band (in this case, over the RF) with more texture of the same type. [The response profile is shown in composite with the response profile for the normal long duration, unchanging homogeneous texture display.] Both the initial texture stimulation and the texture replacement produce direct RF stimulation, in each case resulting in a burst of activity with the normal latency for direct stimulation of approximately 50 milliseconds (arrows).

To study extra-RF modulation, we reposition the RF so that it will fall in the middle of the changing texture band. Now, the RF receives the same static stimulation for the homogeneous texture display and for case where we change texture in the band around the RF. The standard response profile composite for these conditions are shown in Fig. 4B. Here we see that for the texture band change we get no extra-RF modulation. The response to the texture band display does not differ significantly from the homogeneous texture display response, despite the fact that texture replacement outside the RF does activate V1 neurons with RFs in the band itself (as we saw with Fig. 4A). Thus, despite the potent transient in activity that must occur when the texture band appears, this activity is not translated into extra-RF modulation.

Discussion

Given the delay in the expression of extra-RF modulation relative to the activity evoked by direct RF stimulation, we were interested in determining whether this delay constitutes a characteristic feature of extra-RF modulation itself, or whether the delay is somehow contingent on recent stimulation of the RF. The two-step texture stimulation technique that we employed allowed us to distinguish between these possibilities. We found that extra-RF modulation arises with the same timing (and approximately the same amplitude) when the RF and extra-RF areas were stimulated synchronously (the normal one-step case) and when the extra-RF area was manipulated separately later (the two-step case). Thus, extra-RF modulation is not contingent on the burst of activity elicited when the RF of a V1 neuron is suddenly activated. We do not mean to imply that the extra-RF modulation signals we observed would occur without any RF stimulation at all. Rather, our results simply indicate that a V1 neuron may display extra-RF modulation even after its response to RF stimulation has reached a steady-state.

The characteristic delay in the onset of extra-RF modulation that we describe makes it plausible that complex and time-consuming computations may underlay this phenomenon. Complexity in the properties of extra-RF modulation may be reflected in the selectivity ability of the phenomenon to be activated under only certain stimulus conditions (such as, for example, when the RF's home surface appears to contain intrinsic discontinuities.) The two-step texture procedure that we used here allowed us to test the specificity of extra-RF modulation along these lines. While the moat display evoked normal extra-RF modulation, the replacement of texture surrounding the RF with more texture of the same type did not. Although the latter manipulation of the display is a strong stimulus for cells with RFs in the region of texture change, the surface properties of the display do not appear altered as they do with the moat case. Thus, the results are consistent with the surface discontinuity hypothesis, and further bolster the idea that extra-RF modulation reflects sophisticated analysis of visual input.

Figure Legends

1 Response of multi-unit recording site to homogeneous texture display and moat display. The heavy upper line is the response to the moat display. Gray shading illustrates where moat response exceeded homogeneous texture response.

2 Extra-RF modulation with static RF stimulation. **A** shows the response of the multi unit site to a long presentation of the homogeneous texture display. **B** shows the response when the moat is added in the region surrounding the RF 150 ms after the onset of the homogeneous texture display. Extra-RF modultaion appears approximately 100 ms after the appearance of the moat.

3 Comparison of one-step and two-step texture presentation. **A** reproduces the two-step result. **B** reproduces the one-step result. **C** shows the difference between the two-step moat display and the homogeneous texture display. **D** shows the difference between the one-step moat display and the homogeneous display.

4 **A** compares the response to the homogeneous texture display to the case where texture over the RF is replaced by random texture of the same type. **B** shows the result when texture in a band surrounding RF is replaced by random texture of the same type. The band has the same dimensions as the moat. Extra-RF modulation is not evident, although **A** showed that neurons respond to texture replacement over their RFs.

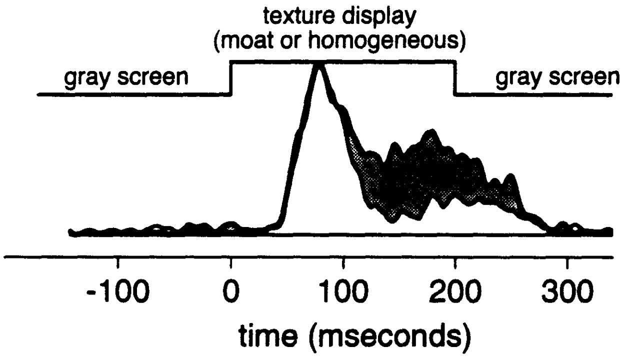

texture display
(moat or homogeneous)
gray screen
gray screen
-100
0
100
200
300
time (mseconds)

Figure 1

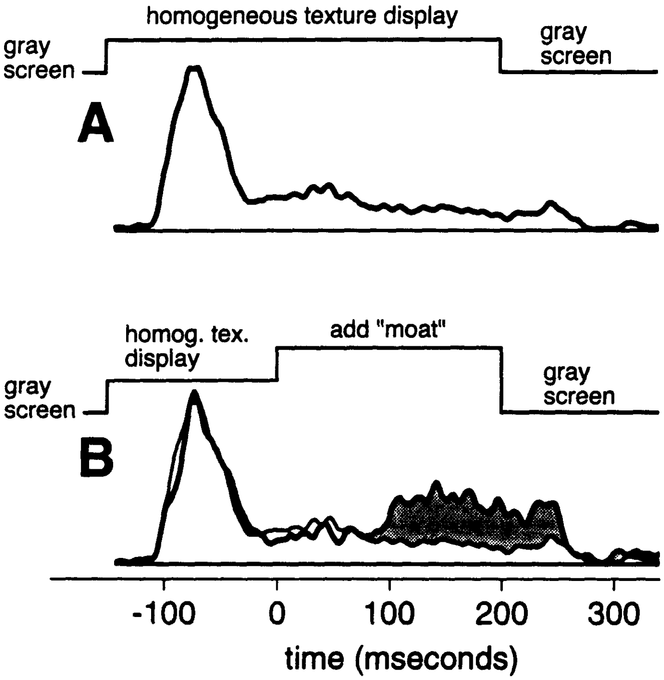

homogeneous texture display
gray
screen
gray
screen
A
homog. tex.
display
add "moat"
gray
screen
gray
screen
B
-100
0
100
200
300
time (mseconds)

Figure 2

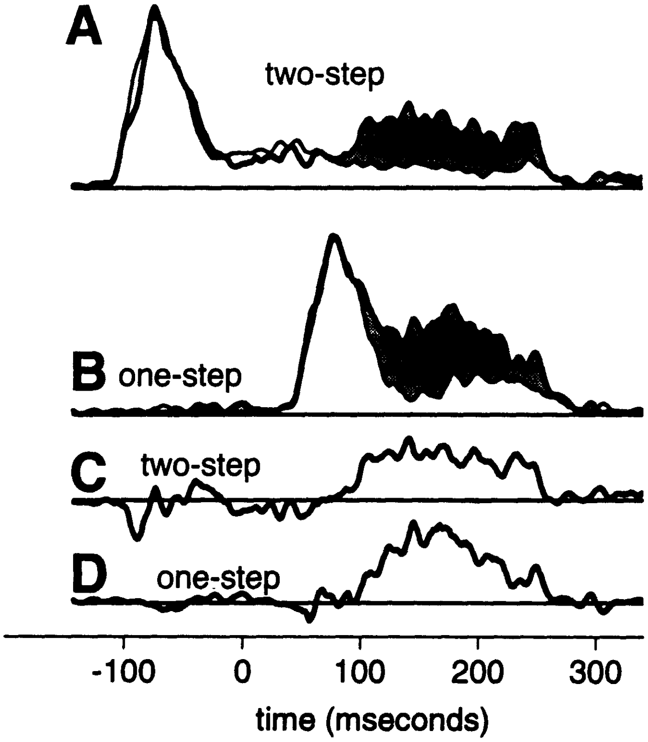

A
two-step
B one-step
C two-step
D one-step
-100
0
100
200
300
time (mseconds)

Figure 3

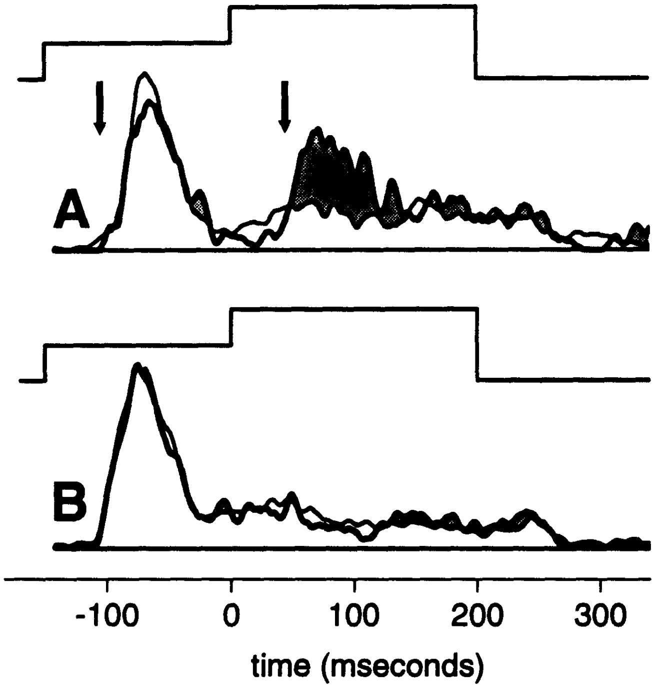


Figure 4

Summary and Conclusion

Summary and Conclusion

Summary of chapter 1

In the first chapter I described experiments designed to investigate basic characteristics of extra-RF modulation among neurons in area V1 of the awake, behaving primate. In the first part of that chapter my colleagues and I asked whether diverse cues may serve to evoke extra-RF modulation in V1 neurons. Of the set of cues that we tested for each cell (i.e., orientation, binocular-disparity, color, and luminance), only orientation had previously been shown to affect extra-RF modulation in primate V1 (Lamme, 1995; Knierim and Van Essen, 1991). Thus, it was certainly conceivable before beginning our experiments that the remaining cues might either not evoke extra-RF modulation at all, or that they might do so in a distinctly different manner for each cue. Instead, we found that V1 neurons do maintain sensitivity for binocular-disparity, color, and luminance cues outside the RF. The extra-RF modulation for these diverse cues was of the same nature as for orientation cues, in that extra-RF modulation for these textured disc displays consistently resulted in responses greater than for the homogeneous texture display.

We found that in some neurons, the modulation evoked by all four diverse extra-RF cues was similar in magnitude. Remarkably, this property of cue-invariance even held for these cells when discs segmented from their background texture by a combination of all four of these diverse cues simultaneously. Approximately one-eighth of cells in the sample displayed this form of cue invariant extra-RF modulation. The remaining cells showed extra-RF cue sensitivity intermediate between invariance and a moderate selectivity for particular cues. Convergence of diverse cues onto individual V1 neurons is a dominant characteristic of extra-RF modulation. This means that cues traditionally considered separate subjects of study, such as color and binocular-disparity, are linked in the sense that extra-RF modulation commonly uses them both. Although it has been suggested that different visual cues, such as color and binocular disparity, are processed independently by separate anatomical modules in the visual system (Livingstone and

Hubel, 1988), our results show that many V1 neurons treat these cues interchangeably, at least from the perspective of extra-RF modulation.

In the second part of the first chapter we provided the first measure of the spatial tuning of extra-RF modulation in area V1 through recordings of individual or clusters of neurons. These data showed extra-RF modulation to operate over a far greater range than previously known (approximately 10° diameter on average for cells with eccentricity between 2° and 6° of visual angle). Taken together, our results support a novel view of V1 in which individual neurons co-process information from diverse cues over a spatial extent very large compared to their RFs.

Implications of chapter 1

The data that I described in the first chapter present us with an interesting dilemma. There I showed that the extra-RF modulation expressed by individual V1 neurons is receptive to diverse extra-RF visual cues, and that the spatial extent over which this extra-RF modulation can be evoked is large compared to RF size. The extraordinary convergence of signals onto individual V1 neurons that I described would seem to lend these cells tremendously increased capacity in visual processing, extending their scope of analysis in both the spatial and the visual cue domains. And yet paradoxically, this convergence of signals also raises an important difficulty for our understanding of the functional role of extra-RF modulation in V1.

The paradox stems from the simple observation that the extra-RF modulation which we described would, in the absence of additional characteristics from those discussed above, convey signals of an extremely ambiguous nature. For example, if orientation and depth cues evoke the same extra-RF modulation in a given V1 neuron, the cell cannot specifically signal which of the two cues was the source of this modulation. The problem is further compounded by the large spatial scope of extra-RF modulation, which reduces the possibility of localizing in visual space the patterns that evoked the modulation.

Although in our experiments this extra-RF modulation distinguished between disc displays and the homogeneous texture display, it is difficult to imagine that a mechanism with the ambiguities discussed above could be of much use in processing natural scenes. In natural scenes, visual cues may be scattered in unpredictable configurations that may or may not delineate well circumscribed figures such as our textured discs. If extra-RF modulation were in this manner receptive to *any* type of pattern heterogeneity in a large region around the RF, it seems unlikely that the mechanism would be of value in visual analysis.

Extra-RF modulation in V1 must then, it seems clear, have some specificity of function beyond what we have described. What might be the nature of this specificity? Previously, Lamme (1995) suggested that extra-RF modulation expresses a specificity of function tied, not to particular visual cues, but to *perceptual* qualities of "figure" and "ground". His evidence for this assertion came from the observation that extra-RF modulation of V1 neural responses was equivalent at different positions across the extent of a moderately sized textured "figure" (like our textured discs), but fell off immediately outside the "figure". Lamme interpreted the responses of the cells he studied as specific signals of whether their RFs were positioned inside or outside of the "figure" region. A simple mechanism of lateral interactions, he argued, would not act in this way, because such a mechanism would have been influenced non-specifically by the presence of the "figure" outside the RF even when the RF itself was positioned on the textured "ground".

Lamme's particular "figure/ground" hypothesis is of course highly speculative. But his broader concept, that V1 extra-RF modulation is linked with *perceptual* interpretations of an image, seems to us very attractive. This is because this concept provides an explanation of why extra-RF modulation maintains the tremendous (and as we discussed above, baffling) convergence of signals from diverse cues across space. The invariances of extra-RF modulation that we observed over visual cues, for example, could reflect the perceptual equivalence of these cues for segmenting the textured disc surface (or "figure") from its background. The large spatial range of extra-RF modulation could grant it sufficient scope to evaluate perceptual qualities of visual surfaces (or "figures") that cannot be appreciated through a purely local analysis.

Is this conceptual link between perception and extra-RF modulation indeed reflected by neural processing in area V1? The key to this question lies in investigating whether extra-RF modulation maintains specificity for perceived image structure. By *perceived image structure* we refer broadly to the organization of parts of a scene into perceptually relevant components such as visual surfaces or "figure" versus "ground", levels of description above basic visual cues such as orientation or color. Lamme's finding of a qualitative asymmetry in extra-RF modulation when a V1 neuron's RF is inside rather than outside a "figure" region of an image provides the first indication that this neural processing relates to perceived image structure. In the second chapter, I described further experiments in area V1 of awake, behaving macaque monkeys that allow us to investigate this question.

Summary of chapter 2

In this chapter I demonstrated the specificity of extra-RF modulation by presenting examples of stimuli that consistently did not evoke this effect. The moat and frame displays allowed us to explore this. These displays featured a rather narrow band of pattern heterogeneity surrounding the RF. However, the power of these displays lies in the fact that they may or may not evoke normal extra-RF modulation, depending on the depth relationships in the display. The moat display, in which the moat appeared to circumscribe the small surface on which the RF was positioned, evoked normal extra-RF modulation, whereas the frame display, in which the frame merely appeared to occlude the large background surface, did not. Here then was a clear example of specificity in extra-RF modulation. We next turned to the question of whether or not this could be accounted for by a simple mechanism.

A very simple explanation for the asymmetry in effect of moat versus frame displays could occur if extra-RF modulation were simply insensitive to the disparity difference between the frame and the background texture, perhaps in analogy to stereo-anomalous subjects that lack sensitivity specifically for crossed disparities (Richards, 1971). We ruled out this hypothesis by comparing the results of the frame & orientation display with the results form the orientation-band display. While the orientation-band display could evoke a modest augmentation in response compared to the homogeneous texture display, this extra-RF modulation went away for the frame & orientation case. If frame disparity were merely ignored, then the augmentation in response should have remained. The results therefore indicated that frame disparity was indeed registered by the mechanisms underlying extra-RF modulation.

At this point a second simple explanation for the moat versus frame asymmetry comes into play. This is the idea that the relative *near* disparity of the frame has a suppressive effect through extra-RF modulation, whereas the relative *far* disparity of the moat has the normal facilitory effect through extra-RF modulation. We countered this idea with the disparity-hole display, the case where normal extra-RF modulation is evoked by displays with relatively *near* disparities surrounding the RF.

Finally we examined the ring display, in which the ring appears to occlude the homogeneous texture surface. We found that this display did not evoke consistent extra-RF modulation in the way that the textured discs from the previous paper did. This is consistent with the hypothesis that cells follow our perception that the texture within the ring actually belongs to the homogeneous texture background.

Implications of chapter 2

In chapter 2 we examined cases where heterogeneity in the pattern surrounding the RF did not evoke extra-RF modulation. As described above, we found that we could explain these results with the hypothesis that extra-RF modulation serves to signal information about whether the RF of a given neuron was positioned on a continuous surface or on a surface with nearby discontinuities, modulation being stronger in the latter case. Of course, this conjecture remains highly speculative. It is certainly conceivable that extra-RF modulation consists of an ensemble of relatively simple mechanisms that perform tasks that merely *appear* to reflect visual surface representation from the limited number of tests we have performed. Future experiments will be needed to distinguish between these possibilities.

Nonetheless, it is interesting to speculate as to the implications of the surface hypothesis. A particularly interesting point relates to the function of extra-striate areas. The standard notion in the past is that extra-striate RFs are built up of the RF characteristics of cells at earlier stages. However, if neurons in area V1 display interesting coding of visual surfaces through extra-RF modulation, it is conceivable that extra-striate neurons (say, in V2) distill this specific information from area V1. Thus, extra-striate cells that in the past have been seen as specific for certain types of bar stimuli (or stimuli of other shapes) may in fact be attempting to coding information about the shapes of *surfaces*.

Summary of chapter 3

The goal of chapter 3 was to determine whether the delay in the onset of extra-RF modulation relative to the initial response of a V1 neuron to direct RF stimulation reflects a central characteristic of extra-RF modulation, or if it rather reflects some artifact of our stimulation technique. The two-step texture stimulation technique that we employed allowed us to distinguish between these possibilities. We found that extra-RF modulation arises with the same timing (and approximately the same amplitude) when the RF and extra-RF areas were stimulated synchronously (the normal one-step case) and when the extra-RF area was manipulated separately later (the two-step case). Thus, extra-RF modulation is not contingent on the burst of activity elicited when the RF of a V1 neuron is suddenly activated. We do not mean to imply that the extra-RF modulation signals we observed would occur without any RF stimulation at all. Rather, our results simply indicate that a V1 neuron may display extra-RF modulation even after its response to RF stimulation has reached a steady-state.

Implications of chapter 3

The delay in the onset of extra-RF modulation is interesting for a number of reasons. One reason is because it suggests that extra-RF modulation is truly a distinct neural process from the normal RF functioning of a V1 neuron. For in contrast to the delay in expressing extra-RF modulation, V1 neurons display their tuning specificity for visual stimuli with their first action potential responses to visual stimulation (Celebrini et al, 1993). A second reason why the delay in the onset of extra-RF modulation is interesting is because it raises the possibility that time-consuming and complex computations must occur before this phenomenon is expressed, as compared to the relatively simple filter operations of the V1 RFs themselves. The extra 30 to 50 milliseconds before extra-RF modulation appears is enough time for visual signals to travel up to extra-striate cortical areas and back to area V1. Thus, the temporal delay raises the possibility that extra-RF modulation could reflect the more advanced neural processing of these higher visual areas. A third reason why the delay in the onset of extra-RF modulation is interesting is because it raises a possibility for the visual system to separate signals pertaining just to RF stimulation (found in the initial response phase) from signals that also reflect analysis of the area outside the RF (found in the later response phase.)

Conclusion

The data described in this thesis provides us with new insight into the neural processing in area V1. The results clearly show extra-RF modulation allows V1 neurons to receive an enormous convergence of information. Interestingly, this convergence is not of a completely random or haphazard nature. Rather, extra-RF modulation tends to follow a number of constraints in its dynamic properties, the manner in which it combines visual cues, and in its sensitivity to image structure. These results raise the interesting (though highly speculative) possibility that V1 plays are role in the representation of visual surfaces.

References

Albrecht D.G., De Valois R.L., & Thorell L.G. (1980) Visual cortical neurons: are bars or gratings the optimal stimuli? Science, 207: 88-90.

Allman J., Miezin F., & McGuiness E. (1985) Stimulus specific responses from beyond the classical receptive field: neurophysiological mechanisms for local-global comparisons in visual neurons. Annu Rev Neurosci 8: 407-430.

Barlow H.B. (1953) Summation and inhibition in the frog's retina. Journal of Physiology, 119: 69-88.

Barlow H.B. (1972) Single units and sensation: a neuron doctrine for perceptual psychology? Perception 1: 371-394.

Bonds A.B. (1989) Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. Visual Neuroscience, 6:239-255.

Celebrini S., Thorpe S., Trotter Y., & Imbert M. (1993) Dynamics of orientation coding in area V1 of the awake primate. Visual Neuroscience 10:811-825.

Connolly M., & Van Essen D. (1984) The representation of the visual field in parvocellular and magnocellular layers of the lateral geniculate nucleus in the macaque monkey. Journal of Comparative Neurology, 226: 544-564.

Derrington A.M., Krauskopf J., & Lennie P. (1984) Chromatic mechanisms in lateral geniculate nucleus of macaque. Journal of Physiology, 357: 241-265.

Desimone R., Schein S.J., Moran J., & Ungerleider L.G. (1985) Contour, color, and shape analysis beyond the striate cortex. Vision Research, 25: 441-452.

De Valois K.K., De Valois R.L., & Yund E.W. (1979) Responses of striate cortex cells to grating and checkerboard patterns. J. Physiol. (Lond.), 291: 483-505.

De Valois R.L., & De Valois K.K. (1990). Spatial Vision. Oxford: Oxford University Press.

Dow B.M., Snyder R.G., Vautin R.G., & Bauer R. (1981). Magnification factor and receptive field size in foveal striate cortex of the monkey. Exp. Brain Res., 44, 213-228.

Enns J.T., & Rensink R.A. (1994) An object completion process in early vision (in press).

Enroth-Cugell C., & Robson J.G. (1966) The contrast sensitivity of retinal ganglion cells of the cat. Journal of Physiology, 187: 517-522.

Freeman R.D, & Ohzawa I. (1990) On the neurophysiological organization of binocular vision. Vision Research, 30(11): 1661-1676.

Gilbert C.D., & Wiesel T.N. (1990) The influence of contextual stimuli on the orientation selectivity of cells in the primary visual cortex of the cat. Vision Res 30: 1689-1701.

Hubel D.H., & Wiesel T.N. (1968) Receptive fields and functinoal architecture of monkey striate cortex. J. Physiol. (Lond.), 195, 215-243.

Hubel D.H. & Wiesel, T.N. (1974) Uniformity of monkey striate cortex: a parallel relationship between field size, scatter and magnification factor. J. Comp. Neurol., 158: 295-306.

Kanisza G. (1979) Organization in Vision. Essays in Gestalt Perception. (New York: Praeger).

Kaplan E., & Shapley R.M. (1982) X and Y cells in the lateral geniculate nucleus of macaque monkeys. Journal of Physiology, 330: 125-143.

Knierim J.J., & Van Essen D.C. (1992) Neural responses to static texture patterns in area V1 of the alert macaque monkey. J. Neurophysiol 67: 961-980.

Kuffler S.W. (1953) Discharge patterns and functional organization of mammalian retina. Journal of Neurophysiology, 16: 37-68.

Lamme V.A.F. (1995) The neurophysiology of figure-ground segregation in primary visual cortex. J. Neuroscience, 15(2):1605-1615.

Lennie P., Krauskopf J., & Sclar G. (1990) Chromatic mechanisms in striate cortex of macaque. J. Neuroscience, 10(2):649-669.

Livingstone M., & Hubel D. (1988) Segregation of form, color, movement, and depth: anatomy, physiology, and perception. Science 240:740-749.

Maffei L., & Fiorentini A. (1976) The unresponsive regions of visual cortical receptive fields. Vision Res 16: 1131-1139.

Marr D. (1982) Vision (San Francisco, CA: W.H. Freeman).

McGuire B.A., Gilbert C.D., Rivlin, P.K., & Wiesel, T.N. (1991) Targets of horizontal connections in macaque primary visual cortex. 305:370-392.

Movshon J.A., Adelson E.H., Gizzi M.S., & Newsome W.T. (1985) The analysis of moving visual patterns. In C. Chagas, R. Gattass, C. Gross (Eds.), Pattern recognition mechanisms (Pontificiae Academiae Scientiarum Scripta Varia, 54) (pp. 117-151). Rome: Vatican Press.

Movshon J.A., Thompson I.D., & Tolhurst D.J. (1978) Spatial and temporal contrast sensitivity of neurons in areas 17 and 18 of cat's visual cortex. J. Physiol. (Lond.), 283: 101-120.

Nakayama K., Shimojo S., & Silverman G.H. (1989) Stereoscopic depth: its relation to image segmentation, grouping, and the recognition of occluded objects. Perception, 18: 55-68

Nakayama K., & Shimojo S. (1992) Experiencing and perceiving visual surfaces. Science, 257: 1357-1363.

Nelson J.I., & Frost B.J. (1978) Orientation selective inhibition from beyond the classical visual receptive field. Brain Res 139: 359-365.

Poggio G.F., & Fischer B. (1977) Binocular interaction and depth sensitivity in striate and prestriate cortex of behaving rhesus monkey. J. Neurophysiology, 40:1392-1405.

Poggio G.F., Gonzalez F., & Friedemann K. (1988) Stereoscopic mechanisms in monkey visual cortex: binocular correlation and disparity selectivity. J. Neuroscience 8(12): 4531-4550.

Robinson D.A. (1963) A method of measuring eye movement using a scleral search coil in a magnetic field. IEEE Trans Biodmed Electron 101:131.

Rockland K. and & Lund J.S. (1983) Intrinsic laminar lattice connections in primate visual cortex. J. Comp. Neurol. 216: 303-318.

Schiller P.H., Finlay B.L., & Volman S.F. (1976) Quantitative studies of single cell properties in monkey striate cortex. I-III. J Neurophysiol 39: 1288-1374.

Ts'o D.Y, & Gilbert C.D. (1988) The organization of chromatic and spatial interactions in the primate striate cortex. J. Neuroscience, 8(5): 1712-1727

Tyler C.W. (1983) Sensory processing of binocular disparity, in Vergence Eye Movements: Basic and Clinical Aspects (Boston, MA: Schor C.M., Ciuffreda K.J.).

von der Heydt R., & Peterhans E. (1989) Mechanisms of contour perception in monkey visual cortex. I. Lines of pattern discontinuity. Journal of Neuroscience 9: 1731-1748.

Wilson H.R., & Humanski R. (1993) Spatial frequency adaptation and contrast gain control. Vision Research, 33:8 1133-1149.