Encoding Surfaces from Motion in the Primate Visual System

by

Stefan Treue
Vordiplom, Biology, Heidelberg University, FRG 1986

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 1992

Signature of Author _____

Department of Brain and Cognitive Sciences
July 30, 1992

Certified by _____

Richard A. Andersen
Professor, Department of Brain and Cognitive Sciences
Thesis Supervisor

Accepted by _____

Emilio Bizzi
Department Head

To my parents

Karl-Heinz & Silke

Encoding Surfaces from Motion in the Primate Visual System

by

Stefan Treue

Submitted to the Department of Brain and Cognitive Sciences on July 30, 1992 in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Neuroscience

Abstract

Moving dynamic random dot patterns were used to investigate the perception and neural processing of structure from motion as well as motion transparency in man and awake behaving macaque monkey.

The main findings were:

(1) Structure from motion perception is based on velocity measurement of object features rather than on successive position measurements. These velocity measurements are spatially and temporally integrated through a process of surface interpolation (Chapter 2).

(2) The interpolated surfaces are used not only as an aide in determining initial depth to new object features but rather is the way in which the object is neurally represented (Chapter 3).

(3) There seem to be neurons in area MT of the macaque monkey that are tuned to the shape of velocity gradients similar to the ones present in our structure from motion displays possible serving as the neural substrate of structure from motion perception (Chapter 4).

(4) Random dot patterns moving across each other in transparent motion are a valuable tool for investigating aspects of motion processing in the primate visual motion pathway (Chapter 5 + 6).

(5) Direction selective cells in area V1 of the macaque monkey as a population seem to segment transparently moving random dot pattern into their component surfaces by ignoring all but their preferred direction. This suggests a mechanism of direction selectivity that is primarily relying on excitatory rather then inhibitory mechanisms. Direction selective cells in area MT of the macaque monkey on the other hand are strongly inhibited by adding their non-preferred direction to a random dot pattern moving in the cells' preferred direction. This suggest an additional stage in cortical processing of visual motion similar to the opponent stage proposed in the Adelson-Bergen model of motion processing in which channels tuned to opposite directions are inhibiting each other possibly to reduce noise in the image. This stage could also serve as a step towards extracting surfaces in motion through spatial pooling of the image data (Chapter 5).

(6) The broad direction tuning curves common to area V1 and MT are sufficient to account for the abilities of humans to discriminate small changes in directions of motion. The smallest discriminanda for individual cells occur not at their preferred direction of motion but rather on the flank of the direction tuning curves. In both area V1 and MT the variance associated with the cells' responses is about equal to the mean response indicating little difference between striate and extrastriate cortex (Chapter 6).

(7) The opponency for direction of motion in the center of MT receptive fields seems to be complemented by an opponency of reversed polarity in the surround, suggesting the existence of 'double-opponent' cells in the awake macaque monkey (Chapter 7).

Thesis Supervisor: Dr. Richard A. Andersen

Title: Professor of Neuroscience

Acknowledgements

I thank Richard Andersen for his support and guidance over the last five years.

I would like to thank all of the past and present members of the Andersen laboratory for providing the fun and stimulation that kept me going through the ups and downs of graduate studies. Special thanks go to Catherine Cooper, Richard Keough, Steve Marchetti and Gail Robertson for outstanding technical assistance and to Ning Qian for his patience in sharing the experimental setup with me and in discussing the intricacies of computational vision.

Special thanks also go to my two main collaborators, Masud Husain and Bob Snowden who not only made the hours of observing rotating cylinders and listening to the crackle of amplifiers so much more bearable but also became dear friends.

I thank Ellen Hildreth, Peter Schiller and Ken Nakayama for serving on my thesis committee and Jan Ellertsen for piloting me through the treacherous waters of graduate school rules and regulations.

I should not forget to mention Irving Diamond and Mike Conley who introduced me to American neuroscience and prevented me from slipping back into the world of molecular biology.

I am grateful for Evangelisches Studienwerk Villigst e.V., Educational Foundation of America and Poitras Fund fellowships which bore most of the financial burdens of the last years.

And last but certainly not least I thank my family and Suzie for their love and support.

Table of Contents

Chapter 1

Preface & Summary

Preface

This dissertation comprises four articles that are either already published (Chapters 2, 5, & 6) or to be submitted (Chapter 3) as well as two chapters containing an outlook on future and ongoing experiments (Chapters 4 & 7). Each chapter is intended to stand alone and therefore all pertinent references and figures are assembled at the end of each chapter. The hurried reader is referred to the second part of this chapter for abstracts of each chapter.

Here I provide an outline of the themes connecting the various chapters. The overarching theme addressed in this thesis is the role surfaces play in visual motion processing in man and monkey. The three chapters in the first part deal with the perception of structure from motion. This refers to the ability to determine the three-dimensional shape of surfaces of moving objects from their two-dimensional projections onto the retina. As my work demonstrates this process involves spatial and temporal integration that is achieved through a process of surface interpolation. An interesting perceptual demonstration described in Chapter 3 strongly suggests that the role of interpolated surfaces in structure from motion goes beyond being just a way to implement the spatial and temporal integration process. Rather the interpolated surfaces seem to be the way the visual system internally represents the observed objects.

I demonstrate that the measurements used in human structure from motion perception are the velocities of object features. The surface interpolation process is therefore based on the velocity flow fields present in the image. This suggests that the visual cortex contains neurons specifically designed to analyze velocity gradients, which are the basis of 3-D shape perception. In chapter 4 I try to establish the presence of such cells in area MT of the awake behaving monkey.

But surfaces play a role not only in the analysis of structure from motion. The findings presented in chapter 5 demonstrate how neurons in primary visual cortex, area V1 of the awake behaving monkey segment transparent surfaces moving across each other into its two component surfaces. In a further step these different directions of motion then interact in area MT in a way that seems to be designed to reduce noise and spatially integrate the motion signals to represent surfaces from motion. Noise in the neural signal is further investigated in the next chapter. Chapter 6 also looks at the issue of how broadly directional tuned cells can contribute to the fine direction discrimination exhibited by human subjects. Finally Chapter 7 returns to the question of how opponency in receptive field organization can be used to segment the visual input into surfaces defined by different directions of motion and/or by surfaces separated in depth.

Abstracts

Chapter 2

Novel dynamic random-dot displays representing a rotating cylinder or a noise-field were used to investigate the perception of structure from motion (SFM) in humans. The finite lifetimes of the points allowed the study of spatiotemporal characteristics with smoothly moving stimuli. In one set of experiments subjects had to detect the change from the unstructured motion to the appearance of the cylinder in a reaction time task. In another set of experiments subjects had to distinguish these two stimuli in a two alternative forced choice task.

The two major findings were: (1) A relatively constant point lifetime threshold (50-85 ms) for perceiving structure from motion. This threshold is similar to the threshold for estimating velocity and suggests that velocity measurements are used to process SFM. (2) Long reaction times for detecting structure (~ 1 sec). The buildup of performance with time and with increasing numbers of points reflects a process of temporal and spatial integration. We propose that this integration is achieved through the generation of a surface representation of the object. Information from single features on the object appears to be used to interpolate a surface between these local measurements allowing the system to improve perception over extended periods of time even though each feature is present only briefly. Selective masking of the stimulus produced characteristic impairments which suggest that both velocity measurements and surface interpolation are global processes.

Chapter 3

Dynamic random-dot displays representing a rotating cylinder were used to investigate surface interpolation in the perception of structure from motion (SFM) in humans. Surface interpolation refers to a process in which a complete surface in depth is reconstructed from the object depth values extracted at the stimulus features.

In one set of experiments trying to establish the spatial extent of the interpolation process, subjects had to detect the presence of a featureless area on the cylinder. In another set of experiments subjects were presented with a variation on the rotating cylinder in which all dots were oscillating either in synchrony or asynchronously. Finally a variety of previously documented perceptual peculiarities in the perception of structure from motion are discussed in light of our surface interpolation hypothesis.

The two major findings were: (1) Featureless areas can be as wide as one quarter of the stimulus before they are reliably detected. (2) Subjects perceive a rigidly rotating cylinder even when such a percept is not in agreement with the physical stimulus. This discrepancy can be reconciled if points that reverse their direction of motion lose their identity and are treated as newly appeared points.

Both of these findings offer strong perceptual evidence for a process of surface interpolation and are also physiologically plausible given results from recordings in awake behaving monkey cortical areas V1 and MT. The companion paper demonstrates how such a surface interpolation process can be incorporated into a structure from motion algorithm and how object boundaries can influence the perception of structure from motion as has been demonstrated before and in this paper.

Chapter 4

This chapter investigates if area MT of the awake behaving monkey contains neurons tuned to velocity gradients. A cell is considered tuned to a velocity gradient if this gradient evokes a response larger than the one elicited by a flat velocity profile moving in the cell's preferred direction at the cell's preferred speed. Moving random dot patterns were used to generate accelerating and decelerating velocity gradients. Eight different gradients of two different slopes each were used. These gradients can be represented in a deformation space in which each stimulus is characterized by a vector. The direction of the vector represents the direction of steepest velocity slope in the stimulus. Preliminary data suggest that area MT does indeed contain neurons tuned to these stimuli.

Chapter 5

An important use of motion information is to segment a complex visual scene into surfaces and objects. Transparent motions present a particularly difficult problem for segmentation since more than one velocity vector occurs at each local region in the image, and current machine vision systems fail in these circumstances. The fact that motion transparency is prevalent in natural scenes, and yet artificial systems display an inability to analyse it, suggests that the primate visual system has developed specialized methods for perceiving transparent motion. Also, the currently prevalent model of physiological mechanisms for motion direction selectivity employs inhibitory interactions

between neurons; such interactions would silence neurons under transparent conditions and render the visual system blind to transparent motion. To examine how the primate visual system solves this transparency problem, we recorded the activity of direction selective cells in the first (area V1) and later (area MT) stage in the cortical motion processing pathway in behaving monkeys. The visual stimuli consisted of random dot patterns forming single moving surfaces, transparent surfaces, and motion discontinuities. We found that area V1 cells responded to their preferred direction of movement even under transparent conditions, whereas area MT cells were suppressed under the transparent condition. These data suggest a simple solution to the transparency problem at the level of area V1. More than one motion vector can be represented at a single retinal location by different subpopulations of neurons tuned to different directions of motion; these subpopulations may represent the early stage for segmenting different, transparent surfaces. The results also suggest that facilitatory mechanisms, which unlike inhibitory interactions are largely unaffected by transparent conditions, play an important role in direction selectivity in area V1. The inhibitory interactions for different motion directions for area MT neurons may contribute to a mechanism for smoothing or averaging the velocity field, computations thought to be necessary for reducing noise and interpolating moving surfaces from sparse information.

Chapter 6

We studied the response of single units to moving random dot patterns in areas V1 and MT of the alert macaque monkey. Most cells could be driven by such patterns; however, many cells in V1 did not give a consistent response but fired at a particular point during stimulus presentation. Thus different dot patterns can produce a markedly different response at any particular time, though the time averaged response is similar. A comparison of the directionality of cells in both V1 and MT using random dot patterns shows the cells of MT to be far more directional. In addition our estimates of the percentage of directional cells in both areas are consistent with previous reports using other stimuli. However, we failed to find a bimodality of directionality in V1 which has been reported in some other studies. The variance associated with response was determined for individual cells. In both areas the variance was found to be approximately equal to the mean response, indicating little difference between extrastriate and striate cortex. These estimates are in broad agreement (though the variance appears a little lower) with those of V1 cells of the anesthetized cat. The response of MT cells was simulated on a computer from the estimates derived from the single unit recordings. While the direction tuning of MT cells is quite wide (mean half-width at half-height approximately 50°) it is shown that the cells can reliably discriminate much smaller changes in direction, and the performance of the cells with the smallest discriminanda were comparable to thresholds measured with human subjects using the same stimuli (approximately 1.1°). Minimum discriminanda for individual cells occurred not at the preferred direction, that is, the peak of their tuning curves, but rather on the steep flanks of their tuning curves. This result

suggests that the cells which mediate the discrimination of motion direction may not be the cells most sensitive to that direction.

Chapter 7

This chapter discusses the possibility that the opponency between directions of motion within the center of the receptive fields of MT neurons as investigated in Chapter 5 is complemented by an opponency of opposite polarity in the area surrounding the 'classical receptive field'. Some preliminary evidence that this is indeed the case for some neurons is presented.

Structure from Motion

Chapter 2

Human Perception of Structure from Motion

Introduction

It has long been appreciated that humans are capable of perceiving the three-dimensional shape of an object using motion cues (Miles, 1931). The shadow of a static, bent paper-clip projected onto a two-dimensional screen, for example, offers little insight into its 3-D shape. Yet, if the clip is rotated so that the shadows of its parts move relative to each other, its 3-D structure becomes immediately apparent (Wallach and O'Connell, 1953) — the so-called kinetic depth effect or "structure from motion" (SFM).

Geometrical considerations show that the recovery of 3-D structure from motion from 2-D images is not a trivial problem. Since there are an infinite number of 3-D interpretations of a given pattern of motion in a 2-D image, the problem is considered to be "ill-posed" (Poggio, 1985). In order to formulate a unique, one-to-one mapping between the 2-D image and the 3-D interpretation constraints are required. Different constraints have been proposed to restrict the range of solutions and have led to the development of a number of computational theories which specify a unique 3-D interpretation for moving elements in a series of 2-D images.

Two general classes of theory have been proposed: those that use velocity information ("velocity algorithms" or "continuous algorithms") and those that employ position measurements ("position algorithms" or "discrete algorithms"). Since both types of theory claim to solve the SFM problem, the question arises as to which, if any, of the several proposed computational algorithms may be employed by the human visual system.

Although several investigators have reported observations regarding the perception of SFM, it has been difficult to use these findings to determine which of the proposed algorithms might be used by the human visual system. Three factors have contributed to this difficulty. First, many of the stimuli used contained non-motion cues, of which the most important is the presence of patterns in the displays which change their 2-D shapes during rotation. Some stimuli used in previous studies of SFM perception contained non-motion cues which allowed subjects to perform above chance in the absence of motion (Braunstein, Hoffman, Shapiro, Andersen, & Bennett, 1987).

The second factor which makes the interpretation of many previous studies difficult is the use of informal observations or subjective assessments of the quality of the 3-D percept. This lack of systematic and objective assessment of human performance might explain, at least in part, why some of the findings have been contradictory (see Dosher, Landy and Sperling, 1989 for review).

The third problem associated with previous studies concerns the investigation of the number of views and features required to perceive SFM. A number of studies have shown that as few as 2-3 different views are sufficient to evoke 3-D percepts (Johansson, 1975; Borjesson and von Hofsten, 1973; Lappin and Fuqua, 1983; Braunstein et al., 1987; Grzywacz, Hildreth, Inada and Adelson, 1988; Todd, Akerstrom, Reichel, Hayes, 1988) but there are contradictory reports on the effect of numbers of points on the saliency of SFM perception. Two groups report that SFM perception is improved with increasing the number of dots (Braunstein, 1962; Sperling, Landy, Dosher and Perkins, 1989) while one group of investigators claims that increasing the number of moving elements has no effect in their task (Todd et al., 1988). Another study contends that performance actually deteriorates (Braunstein et al., 1987). Similarly, although a

larger number of frames appears to be helpful in creating more stable 3-D percepts (Wallach et al., 1953, White and Mueser, 1960, Green, 1961, Doner et al., 1984, Braunstein et al., 1987, Grzywacz et al., 1988), it has also been reported that observers are able to perceive structure from as few as two different frames (Lappin, Doner and Kottas, 1980, Doner, Lappin and Perfetto, 1984, Todd et al., 1988).

Another direction of research has attempted to find similarities between the performance of computational algorithms and recorded psychophysical performance. Grywacz et al. (1988) demonstrated that perception of SFM builds up with increasing extent of angular rotation of a smoothly moving dot display. They interpret this finding as evidence for the use of the incremental rigidity scheme. This elegant position-based algorithm proposed by Ullman (1984) has been shown to perform best when the angular distance travelled by the points between the samplings is large and the stimulus is viewed for an extended period of time. However, three recent studies (Petersik, 1987, Todd et al., 1988, Mather, 1989) show that increasing the angular displacement of points between frames degrades perception of SFM.

In the present study, we use novel dynamic random-dot stimuli developed in our laboratory to investigate the spatiotemporal characteristics of human perception of SFM. We have two aims. The first is to delimit the minimal information required in both the spatial and temporal domains to evoke 3-D percepts from 2-D motion cues and to investigate how performance changes with manipulations of these stimulus parameters. In particular, we are interested in examining how the spatial and temporal factors interact. These interactions may help to explain inconsistencies between previous reports

since many of these have employed only a few combinations of parameters. Several features of our stimulus minimize the problems associated with some previous investigations. The use of finite lifetimes leads to the reduction of shape cues in the display. It also proves a very powerful tool for investigating the temporal aspects of SFM perception and, together with the elimination of density cues from our displays, gives us rigorous control over stimulus parameters. In contrast to many other studies, we use a high display rate (70 Hz) and movies of several seconds length without cycling through the same set of frames. This allows for the more natural impression of continuous motion and prevents the subjects from memorizing the stimulus. Finally, our stimulus allows us to use reaction time and forced choice experimental paradigms to assess perception of SFM quantitatively.

The second aim of our study is to investigate the kind of object representation generated in the SFM process. Investigations in depth perception using disparity information have demonstrated surface interpolation between feature elements (Collett, 1985; Morgan & Watt, 1982; Mitchison & McKee, 1985; Würger & Landy, 1989). Position-based SFM algorithms like Ullman's (1984) incremental rigidity scheme on the other hand generate wire-frame models from the visible features. These algorithms require the continuous presence of all features during the computation. This issue of whether object features have to be continuously present is of some biological significance because under natural viewing conditions objects are often opaque and features rotate out of sight. In this study, we investigate whether surface representations, which do not require the continuous presence of all the points in the display, may be used for the perception of SFM.

General Methods

Stimuli and Protocol

The stimuli were dynamic random dot displays presented on a CRT screen. The dots on the screen were the orthographic projection of points on the surface of a transparent, rotating cylinder. They lived for a pre-determined number of frames (finite lifetime) and appeared and disappeared asynchronously. These "flickering" displays minimize position cues in the stimulus since any configuration of dots will dissolve within a short time. In this way we attempt to examine the responses of the visual system to motion information alone.

If the average dot density is constant over the surface of a cylinder, its projection onto a two-dimensional surface yields an image with a greater density of dots at the edges representing the sides of the cylinder. In order to eliminate this density cue, the random positions of dots are first generated in screen coordinates and then projected orthographically onto the surface of the cylinder. Since we repeat this procedure for every point when it gets replotted at the end of its lifetime there is always approximately equal density at all locations on the cylinder at any time.

The cylinder is rotated about a fixed, vertical axis. The parallel projection of the moving points is the display viewed by the observer on a CRT screen (Fig. 1). The horizontal velocity profile of the stimulus forms a half cycle of a sinusoid between the two sides of the display while the velocity does not change along any vertical line. In the stimulus points therefore speed up while moving towards the middle of the display where the highest speeds are encountered and they slow down when

heading towards the edges of the display.

Frames for the stimuli are computed off-line and entire movies are stored in the memory of the PDP 11-73 computer used for these experiments. The maximum possible length for each movie was 390 frames. The display rate of 70 Hz yielded movie lengths of 5.5 sec. Since the cylinder is transparent, half the points move in one direction and the other half in the opposite direction. The assignment of one direction of motion to the perceived front or rear surface of the cylinder is ambiguous and may reverse perceptually during viewing.

We refer to this cylinder stimulus as the "structured" display. Our "unstructured" stimulus is computed by randomly reshuffling all the vectors (i.e. the paths the dots travel in their lifetime) in the structured display (points falling off the edges of the display are wrapped around on the other side). Thus, the unstructured stimulus is made using the same set of vectors as in the structured case[1]. However, the structure of the velocity field (the distribution of velocity vectors across space) is altered. Perceptually, the unstructured stimulus appears to some observers as visual noise and to others as a cylinder filled with dots (the structured display being seen as a hollow one). Williams and Sekuler (1986) report a similar observation when they constrain the directions present in a random dot pattern in which dots move in different directions.

In some experiments subjects were shown movies which contained a transition from the unstructured to the structured

[1] It should be noted that this is different from the "no correlation" stimulus employed by Newsome and Paré (1988) and Downing and Movshon (1989). Rather than shuffling the motion vectors in the display these researchers introduce noise by randomly repositioning a certain percentage of the points between frames. In contrast to our procedure the averaged motion in their noise stimulus is therefore very different from the averaged motion in the signal stimulus.

stimulus and were asked to detect the appearance of the rotating cylinder in a reaction-time task. In order to avoid artifactual cues, the transition between unstructured and structured displays is not an abrupt one. Operationally a frame number for transition is selected, however, a point always completes its pre-determined trajectory, even if it lives through the frame designated for transition. When such a point "dies" its new path is appropriate for the velocity field describing the structured stimulus. Thus, the transition is a period beginning with the designated frame for transition and completed within the point lifetime designated for the stimulus.

In other experiments subjects were required to discriminate between the unstructured stimulus and the rotating cylinder in a 2 alternative forced choice task. The two alternative forced choice task was employed whenever the stimulus duration was a crucial parameter.

In the reaction time paradigm each block of trials contained one movie which was identical to one of the others except it did not change to the structured stimulus (the "catch" trial). In this way we could gauge the approximate number of trials in which the observers obtained hits without a corresponding change in structure of the display (i.e. the chance hit rate).

Depending upon the difficulty of the task, subjects were given a reaction time (RT) window starting after 300-500 ms and ending after 1500-2000 ms (0 ms referring to the onset of the transition period). If subjects responded within this window they received a short feedback tone — even if the trial was a catch one, since such trials were also allocated a randomly positioned RT window[2]. The change from the unstructured to

2 Since the "hit rate" on catch trials was generally lower than about 10-15% and

the structured stimulus occurred randomly between 1 and 3 seconds after the beginning of the movie. This proved sufficient to keep a low chance hit rate (~10-15%), i.e. the percentage of catch trials in which subjects responded within the RT window. In one set of experiments in which very low number of points were used the period before a change in structure was randomized over 1 to 7.7 seconds. The reason for lengthening the foreperiod for this more difficult task was that subjects guessed more, as indicated by the increased number of hits in the catch trials. To maintain the chance hit rate at 10-15%, the foreperiod was lengthened. The frame rate was lowered to 35 Hz in theses experiments.

The presentation of a movie ended with either the release of the key by the observer or the end of the RT window. A hit reflects the detection of the change from the unstructured to the structured display within the RT window (% hits = # of correct responses / # of correct + # of late responses). The trials being in which the subjects released the key before the RT window were not included in the computation of the hit rate. These early responses were not used in computing the hit rate to allow the use of different foreperiods, since they were more likely to occur in the tasks using longer foreperiod ranges.

The experimental subjects viewed the display binocularly in a dimly lit room from a distance of 57cm. Head movements were not restricted. The display subtended a visual angle of 6 x 6 degree and had a mean luminance of 1 cd m-2. The size of the single points was about 0.5 mm (subtending 3' visual arc) in diameter. A typical run would contain several movies which were presented in a random block design. The computer would randomly select the movies going through the whole set before a

only one movie in a block of about ten was a catch, only in about 1% of all trials the subjects got such a "misleading" feedback.

new cycle would begin. Typically, 100 stimuli were presented within a run, lasting about 10 to 15 min. After each run the movies were discarded and new ones were generated using new random number seeds. Thus the dot patterns were different for each block of trials. This avoided the possibility of subjects memorizing a movie.

The parameters of the stimulus systematically varied in this study were the number of points, the point lifetime, the angular rotation rate of the cylinder, the movie length, and how much of the stimulus was visible to the observer. For the conditions described above the average 2-D velocity of a stimulus with an angular rotation rate of 35 degrees per second is 1.2 degrees per second. For a lifetime of 100 ms this corresponded to a pathlength of 7 minutes of visual angle. Two highly trained subjects (authors MH and ST) with corrected vision and no history of eye movement abnormalities were used. Since the subjects showed very similar performance all results presented here (except Fig. 2) are averages across them to smooth the curves and unclutter the figures. Whenever specific values are referred to in the text which are different for the two subjects their data are presented separated by a slash (/).

Results

The use of limited point lifetimes shows a minimal temporal requirement

The most important feature of our stimulus, especially in comparison to previous work on the perception of SFM, is the finite point lifetime. Our first experiment investigated the influence of this parameter on the performance in a reaction time task. Each block of trials consisted of 10 movies (including the catch stimulus). Each of these movies had a different point lifetime ranging from 42 to 266 ms. The number of points was kept constant at 128; the angular rotation rate was 35 degrees per second.

The results for the two subjects are shown in Fig. 2. The psychometric curves demonstrate a minimal temporal requirement for the perception of SFM. Subjects cannot perform the task below a point lifetime of ~ 60 ms and the threshold (i.e. the point at which the curve reached 50% of its final height) is at ~ 69/81 ms. Peak performance is reached at a point lifetime of ~ 125 ms.

The threshold for the perception of SFM does not simply reflect a threshold of motion perception per se since subjects can clearly perceive the horizontal direction of motion in the stimulus when lifetimes where so short that no reliable percept of the 3-D object could be achieved. In other words, a longer integration time is needed to perceive SFM than just the direction of motion.

Threshold behavior with changes in point number and angular rotation rate

We investigated the influence of stimulus parameters on the threshold by varying the number of points used for our stimuli between 2 and 128 points. The task and all other parameters were similar to the first experiment. Fig. 3 shows the results. There is a marked decrease in performance when the number of points was lower than 32 but even at 4 points both subjects were still able to reach a performance of over 50%. Below saturation, a doubling of the number of points improved peak performance by ~ 16%. The threshold of all subjects' curves was between 75 and 90 ms and changed very little with large variations in the number of points and peak performance.

The next experiment studied the influence of the rotation rate (and therefore velocity in 2-D and 3-D) on performance. Fig. 4a shows the results of our reaction time task. In contrast to the previous experiment all 4 curves have a very similar shape. The threshold was found to decrease somewhat with increasing rotation rate but only by about 40% (from ~ 49/51 ms at 140 deg s-1 to ~81/87 ms at 21 deg s-1) over the 7-fold increase in rotation rate.

To investigate the possibility that improved performance at higher velocities reflects the increase in angular pathlength we plotted performance for the different rotation rates against the angular pathlength of the points. Fig. 4b shows that the threshold for pathlength was far from constant but rather increased by ~ 300% (from 1.7 to 6.5/7.0 deg) with the increase in speed.

An explanation for the threshold we observed comes from

studies investigating optimal temporal properties of motion. We replotted the data from several such investigations reviewed by Nakayama (1985, his Fig. 6) in Fig. 4c together with our point lifetime thresholds (filled circles) and the points where performance peaks in our task (filled squares). Note the good correspondence over the wide range of tasks used in the different studies. We will argue in the Discussion below that these and other results suggest that precise measurements of velocity are important in the perception of SFM.

Investigating buildup of performance

If indeed velocity measurements are employed for the perception of SFM the question arises as to how these measurements are used to compute 3-D shape. As in the previous experiments our use of finite point lifetimes proves to be an important tool to study how the visual system uses velocity measurements to compute SFM.

The current position-based algorithms, most notably Ullman's incremental rigidity scheme, sample the positions of a set of points in several discrete images and measure how the points change their positions relative to each other between these discrete images in order to compute their 3-D positions correctly. For such a scheme to work, all points have to be present during the entire viewing period so that their relative positions can be assessed. Two groups of investigators have already shown that the perception of SFM is possible with limited lifetimes (Todd et al., 1988, Dosher, et al., 1989). But to truly test the biological validity of Ullman's algorithm, one in addition has to show that stimuli exist for which the full SFM percept is not already achieved with just the information from one lifetime.

Fig. 5 shows the reaction times for detecting the change from an unstructured stimulus to the rotating cylinder. They are plotted together with the respective point lifetimes of the stimulus. Two findings become immediately obvious: 1. Reaction time varies as a function of point lifetime, ranging from ~ 1000 ms to ~ 700 ms for the point lifetimes tested. 2. The reaction times were always several times longer than the point lifetimes.

This second point attracted our attention because it suggests that the visual system is able to integrate information carried by points which appeared at different times. Unfortunately, it is not possible to determine definitively from these data how much of the reaction time is comprised of visual input and how much of it is computation time in the brain or motor reaction time. But such a measurement is crucial in light of studies showing that SFM can, under certain circumstances, be perceived with just two frames (Lappin et al., 1980; Landy, Dosher, Sperling and Perkins, 1988). These studies seem to suggest that the brain might only need to view the stimulus for one point lifetime and then needs all the rest of the reaction time to process the information and execute the motor behavior. In order to investigate this issue subjects were shown stimuli of varying duration in a two-alternative forced-choice experiment. They were asked to report whether they saw a cylinder or unstructured stimulus. We varied the stimulus duration between 42 and 1680 ms and presented equal numbers of structured and unstructured stimuli. The lifetime was kept at 100 ms throughout this experiment. The results are plotted in Fig. 6a (desynchronized point lifetimes). Although subjects performed slightly above chance with a stimulus duration of one lifetime, there is a clear buildup in performance with increasing stimulus duration indicating that information is integrated over several

point lifetimes.

One possible argument consistent with position-based approaches is that the visual system attempts to find a set of points whose lifetimes are aligned (in time) so that it can follow their composite pattern for several frames. Since our lifetimes are desynchronized it would be difficult to find such a set, especially since the number points is large. To control for this possibility we ran an additional experiment which was identical to the previous one except that all the point lifetimes were synchronized so that they all began and ended their "lives" together. These results are also plotted in Fig. 6a (synchronized lifetimes). The synchronization of lifetimes had no effect on the buildup of performance.

A way in which the visual system could perform the observed integration of information would be by fitting a surface through the data points (see Discussion for details). Such a surface interpolation scheme may only be used when the density of the points is high and already closely approximates a surface. We therefore repeated our previous experiment with a cylinder composed of just 12 points. Fig. 6b (filled triangles) shows the same phenomenon as the previous experiment with the buildup taking even longer. This result suggests that surface interpolation is also used in low density displays.

We performed another two alternative forced choice experiment to investigate whether the observed buildup is due to factors other than the use of information from points which were widely separated in time. Movies were created with the same parameters as in the immediately preceding experiment except every point, after living through its first lifetime, was not randomly replotted but rather repositioned at its original

location. It then moved through the same path as before, just to be repositioned at the original location, starting the cycle again. These movies contained the same number of points with the same point lifetime as used for the previous experiment but after the passage of the first point lifetime the stimulus contained no new information. Fig. 6b (open squares) plots the results. It is obvious that the subjects are not able to perform the task, no matter how many lifetimes the stimulus was presented. In summary, we take the results from these experiments as strong evidence for the use of surface interpolation in the perception of SFM.

Global process in the perception of SFM

If SFM is perceived by making use of interpolated surfaces it is important to show that the tasks are performed using global rather than local cues. Due to the nature of our stimulus it is conceivable that subjects could solve the tasks by monitoring local velocity coherence even though they were asked to use the perception of shape from motion as the only cue for responding in the task. In order to exclude the use of only local changes in velocity to perceive SFM, we ran a set of control experiments by masking out portions of the display.

In a reaction time task the subject (ST) was presented with movies in which most of the stimulus had been masked (this was achieved by simply not plotting the points falling within the mask) except for a square of 2 by 2 deg of visual angle in the center. The visible area (~ 11% of the cylinder area) contained an average of 14 points. Since the percept of a full cylinder was obviously impossible, the subject was asked to respond to the appearance of two curved surfaces consistent with an interpretation of a partial view of a rotating cylinder. The results are plotted in Fig. 7a in comparison to the performance achieved

using a full size stimulus with either the same number of points or the same dot density. As is apparent from the data in Fig. 7a it is nearly impossible to perform the task when the mask is present.

We performed another set of experiments to investigate the influence of the size, shape and position of the mask. In these experiments only 25% of the stimulus was masked away. Sketches of the different masks are shown in Fig. 7b. The two types of vertical masks cover the edges or the center of the display. The horizontal masks also cover the center or the edges but the areas they occlude are redundant because all velocities are still represented in the unmasked areas of the cylinder. The horizontal masks were included to control for the effect of disrupting the stimulus by breaking it into two parts (central mask) and for the effect of decreasing overall stimulus size (central and peripheral mask). The results are plotted in Fig. 7c. The data for the peripheral vertical mask (filled circles) show a marked reduction in performance in comparison to the horizontal peripheral mask (filled squares). The results using the central vertical mask show an even stronger reduction in performance (open circles). Performance with a horizontal mask was similar to that measured with an unmasked display (open and filled squares compared to dotted line).

General Discussion

In these experiments we have attempted to examine the spatiotemporal characteristics of the perception of SFM. We used a reaction time task in which subjects were asked to detect a change in the structure of the presented stimulus and several two alternative forced choice paradigms.

We showed:

1) There is a point lifetime threshold for detection of SFM which remains fairly constant (50-85 ms) over a wide range of number of points and velocities, although it does increase somewhat with decreasing angular rotation rates. (Fig. 2, 3, 4)

2) This threshold reflects a minimal temporal requirement and not a minimal pathlength (or threshold for detection of motion) (Fig. 4).

3) Reaction times (RTs) for perceiving SFM are long (Fig. 5), presumably reflecting a process which integrates information temporally across several points lifetimes (Fig. 6). This process is global as shown by the effect of even a small mask on performance. (Fig. 7)

Temporal characteristics

One of the most interesting findings of our study is the demonstration of a point lifetime threshold for the detection of SFM. The sharp drop in performance for point lifetimes shorter than 50-85 ms indicates a minimal temporal requirement. Why does the visual system need point lifetimes of at least this duration ? And why does this required time fall with increasing angular rotation rate ?

One explanation might be that in order to compute the 3-D

locations of points, the visual system samples their 2-D positions in several discrete images and measures how these change relative to each other between frames to compute their 3-D locations. The observed threshold might therefore be attributed to a requirement to sample a minimum number of images in order to assign 3-D positions correctly. Ullman (1984) has in fact proposed such a position-based algorithm — the incremental rigidity scheme — which over a number of images correctly assigns 3-D locations to every point. The observed threshold could therefore represent the minimum point lifetime that is needed for the visual system to obtain enough samples of the the positions of the points. Note that for such a scheme, velocity measurements are not required. In fact, Ullman's scheme does not put any restrictions on the sampling frequency or even the order of frames, therefore making even implicit velocity calculations effectively impossible.

The performance of this position-based scheme depends on the accuracy with which the displacements of points between frames is assessed. It is therefore highly sensitive to errors in measuring 2-D positions if the displacements between sampled images are small. As these displacements become smaller, the algorithm requires larger and larger amounts of rotation to approximate the correct solution (for a detailed analysis of this problem see Grzywacz and Hildreth 1987). Indeed, Grzywacz and Hildreth have shown that a continuous implementation of the incremental rigidity scheme is highly unstable. These authors therefore suggest that continuous motion is discretely sampled (to obtain position) and the actual motion between the samples is used only to solve the correspondence problem between points in the images.

The preceding two paragraphs appear to offer a possible

explanation for the minimal temporal requirement as well as its shift with increasing rotation speed: Position-based approaches such as Ullman's need a set of discrete images of a point but the exact number of views might not be critical as long as a minimum overall pathlength (and therefore the overall angular extent of rotation) is inspected. However, Fig. 4b shows that no speed independent pathlength threshold exists. Position based algorithms would have predicted that the extent threshold would be relatively constant since extent of the movement (and not velocity) is the important parameter for these algorithms. Grzywacz and Hildreth's (1987) analysis points to another important problem: Fig. 4b shows that subjects in our experiments reach peak performance with only a few degrees of angular rotation for each point. Even disregarding the finite lifetimes the rotation of the whole cylinder was only about 30 angular degrees before the subjects responded. The analysis presented by Grywacz et al. (1987, see their Fig. 2 and personal communication) shows that the incremental rigidity scheme is unable to perform accurately which such small amounts of rotation. Thus, it seems unlikely that the sampling requirements of a position-based algorithm such as Ullman's can account for the threshold or the high level of performance observed in our experiments. The need to track object features over extended rotation angles points to another problem of algorithms such as Ullman's: they are vulnerable to occlusions and rotation out of sight when objects are opaque.

At this point it might be useful to point out that occlusions can occur in two different domains. In the object domain the occluder hides part of the object (thereby also limiting lifetimes). Our masking is an example. The other type of occlusion occurs at the level of the features; an example would be when a feature on an opaque rotating object rotates out of

sight. Our use of limited lifetimes is an example of feature based occlusion. Our experimental results clearly demonstrate that the visual system is able to cope with feature based occlusions whereas algorithms which have to track features continuously fail.

An alternative means of solving the SFM problem is to use velocity information rather than the raw positions of points in each image. McKee and Welch (1985) showed that subjects need to view a moving bar for 80-100 ms for asymptotic velocity discrimination. Because this range of point lifetimes corresponds very closely to the threshold observed using our SFM stimuli we suggest that the point lifetime thresholds observed in our experiments reflects the time required to measure velocity accurately. Our explanation for the threshold, therefore, is that subjects use velocity measurements to solve the SFM problem and simply need to see points for a minimum time before they can correctly measure their 2-D velocities. This interpretation of our data is strengthened by findings from several laboratories which show that increased velocities allow the same level of performance with shorter stimulus durations (McKee et al., 1985, De Bruyn and Orban, 1988) and that optimal temporal displacements in apparent motion sequences are shorter for faster velocities (Nakayama, 1985). Fig. 4a shows that perception of SFM has very similar characteristics and the correspondence with data from visual motion experiments is further emphasized in Fig. 4c. This shows how the data presented here compares with the spatio-temporal characteristics of velocity perception (from Nakayama, 1985). The requirement for accurate velocity measurement offers an explanation for our observation that stimuli could be perceived as moving at lifetimes too short for reliable SFM perception. The detection of motion alone is therefore simply not sufficient

for perception of SFM; and the somewhat longer lifetime needed to measure velocities accurately is what leads to the difference.

Support for this velocity hypothesis comes from the finding that lesions of area MT, a region in primate visual cortex which contains neurons tuned to stimulus velocity (Maunsell et al., 1983), have been found to impair perception of SFM (Siegel et al., 1986). The results of two recent psychophysical studies are also in agreement with this hypothesis. Mather (1989) has shown that SFM depends on the outputs of low-level or "short-range" motion detectors and Dosher et al. (1989) argue against position-based algorithms because their subjects showed only a weak loss of performance if the point lifetime was reduced from 30 to 2 or 3 frames. This finding might seem surprising in light of the strong effects of point lifetime in this range in our study but two factors might explain this discrepancy. First, Dosher et al. use very long stimulus onset asynchronies (SOAs) and as a result their shortest point lifetimes were 133 and 200 ms. Their slight decline in performance for two frames might therefore represent the high end of our SFM threshold. Second, it is not so surprising that performance improves so little from 2 to 30 frame lifetimes in light of a recent study by Snowden and Braddick (1989). These investigators showed that long SOAs severely inhibit the ability to use temporal recruitment (Nakayama and Silverman, 1984) to improve performance in long frame sequences as compared to short frame sequences in a direction of motion task. This effect is so pronounced for SOAs between 50 and 100 ms that performance peaks after only about four displacements. These data suggest that Dosher et al. observed such a small improvement with longer lifetimes (i.e. larger frame sequences) because their choice of SOAs effectively prevents the visual system from making use of the more numerous frame sequences. Also they used conditions in which

performance was high even for short lifetimes and this may also account for why they saw so little improvement.

Surface interpolation and spatiotemporal integration

How might velocity measurements be used to compute 3-D shape ? Several algorithms which require the measurement of velocity (or higher derivatives) have been proposed (Clocksin, 1980, Longuet-Higgins & Prazdny, 1980, Hoffman, 1982, Koenderink and van Doorn, 1986). Recently, we suggested that a means of solving the SFM problem is to measure the velocities of as many points as possible across the surface of the object, fit a smooth 2-D velocity field to the measurements, and use this velocity map to derive a 3-D surface representation of the object (Husain, Treue & Andersen, 1989). In theory, such a representation may also be computed by fitting a smooth surface through the 3-D positions derived from the 2-D velocities of each point. These 3-D positions could only be assigned after comparing velocities across the stimulus since depth is determined by relative velocities between different parts of the stimulus rather than absolute local velocities. If such a global comparative process is already performed at the 2-D level it seems more parsimonious that the smoothing and the interpolation of the velocity field is also done at the same time. On the other hand the process of surface interpolation might operate on the 3-D positions to allow for cue-integration especially for stereopsis. Surface interpolation offers a more plausible solution to the SFM problem than current position-based algorithms given that the visual system evolved in an environment where tracked individual features frequently are only present for a short period of time.

In the experiments reported here, we found that RTs for

detecting the change from unstructured to structured stimuli were remarkably long, ranging from 700 to more than 1100 ms. These values are much longer than the point lifetimes needed for perceiving SFM and suggest that the visual system samples the stimulus for several times longer than the lifetime of any one set of points. This observation is consistent with the scheme outlined above: the system integrates measurements from several sets of points to compute a reasonably accurate surface representation.

Our two-alternative forced-choice experiment demonstrates that the long RTs reflect visual input of several lifetimes and not just a long computation or motor response time (Fig. 6b). Such behavior does not exclude the possibility that a position-based algorithm may be employed, but clearly, it cannot be explained by algorithms such as Ullman's incremental rigidity scheme. Furthermore, the results support the surface interpolation hypothesis described above. The possibility that the observed buildup in performance is not due to actual integration of information over time but might just reflect an unrelated intrinsic phenomenon of the visual system is countered by our control experiment in which points were replotted so that they moved through the same paths over and over again. Performance did not increase above chance (Fig. 6b).

There is no a priori reason why the visual system should not improve its performance when repeatedly presented with the same set of frames. It may be that subjects are not able to use fully the information presented in the first few frames, especially if the stimulus appears on an otherwise dark and featureless screen. This is especially true, if the number of different frames is small and the motion measurement across space is not totally in parallel but involves some "patch by patch" measurements. It

has previously been observed "that the perception of rigid rotation from two-frame sequences may be critically dependent on a repetitive oscillation of the display" when high-density stimuli were used (Todd et al., 1988). In our displays, when 12 points were made to retrace their steps over and over again, they probably carry so little information that it can be measured within the first lifetime. In agreement with these considerations we observe some buildup in performance when using high number of points. Under better conditions (i.e. if long point lifetimes, high number of points, and higher rotation rates are used) good performance can already be achieved within the first point lifetime (as also seen by Landy et al.,1988).

It could be argued that our task, since it does not require the subjects to distinguish between two different structures, can be solved by just measuring local coherence of the velocity field. This seems unlikely given that our subjects were asked to use the overall shape of the stimulus as the cue for their response. But stronger experimental evidence comes from our masking experiment. The results show that occluding the cylinder by as little as 25% leads to a marked reduction in performance (Fig. 7a). This is only the case if the mask covers non-redundant areas of the cylinder. If the same size mask is placed horizontally performance is not reduced. This control rules out effects of the number of dots or their density. If subjects monitor coherence locally the observed effect would not be expected since they could easily shift their attention to any unmasked region of the cylinder. The observed reduction of performance with the peripheral mask cannot be due to the subject's preferred monitoring of the stimulus edges since the central vertical mask leads to an even stronger reduction in performance. This result would be expected with a surface interpolation algorithm since it would have problems

interpolating across the central mask while at the edges where the velocities are already low the velocity field would be smoothly interpolated to the stationary surround.

Our use of limited lifetimes turns out to be critical in ruling out another possible explanation of the above results. Would we have used unlimited lifetimes one could have argued that the performance with the vertical mask was degraded because individual point paths are cut short by the mask and therefore are less visible and pool-able. But given that the vertical peripheral mask leaves more than 120 angular degrees of each surface of the cylinder visible and the points travel only through about 2-7 angular degrees (depending on the lifetime used) the above argument does not apply.

Spatial characteristics

The surface scheme depends critically upon integrating over space and time samples taken at different positions across the surface of the object. It predicts that performance should improve as the number of samples taken per unit time increases. In agreement with this prediction, we found that (for the range 2-32 points) peak performance improved with increasing number of points (Fig. 3).

Our results are not in agreement with those of Braunstein et al. (1987) who showed that increasing number of points between 2 and 5 actually worsened perception of SFM. The most likely reason for this discrepancy lies in the differences between the two tasks. Braunstein et al. noted that: "The theoretical analyses considered in the present study were concerned with recovering depth coordinates for individual points. The task that we used would not be appropriate for studying analyses concerned with recovering surface structure. Different results

might be expected for number of points if the task involved detection of surfaces or discrimination among surfaces." It is of interest to note that Braunstein and his colleagues report that their subjects were able to see structure even when just one frame of movie was displayed, suggesting that pattern information as well as motion was available as a cue in their displays.

Our data demonstrate clearly, as others have done previously, that it is possible to see structure with some degree of reliability with only a few points. This is to be expected from the surface scheme outlined above since, provided the viewing time is long enough, the spatial sampling will be sufficiently dense (due to temporal integration) to compute a surface representation.[3] This suggestion of a trade-off between viewing time and number of points is strengthened by our results presented in Fig. 6a and 6b. These data show that the visual system integrates over a longer period of time before it reaches peak performance if fewer points are presented. What is not expected a priori from the surface hypothesis, however, is the finding that peak performance fails to reach 100% correct responses when low numbers of points (2-16) are used, since in principle, there should be sufficient data present to extract the 3-D structure given a long enough stimulus duration. This result therefore suggests a limited spatio-temporal memory capacity of the SFM system.

[3] This is also our explanation for the ability to perceive at least a very crude perception of SFM with as little as 2 points.
Interestingly the percept of such a sparse stimulus at the short lifetimes used in our experiment is one containing about 5-10 rather than just two points which is additional perceptual evidence for temporal integration.

Algorithms and Motion Transparency

The proposal that the visual system integrates many samples over space and time to compute a 3-D surface representation of the object has also been advanced to account for perception of short range coherent motion (Van Doorn and Koenderink, 1984, Snowden et al., 1989). A large number of algorithms for 2-D velocity measurement have been proposed which perform some velocity integration, averaging or smoothing (Hildreth and Koch, 1987, Horn and Schunk, 1981, Zucker and Iverson, 1986, Yuille and Grzywacz, 1988, Bülthoff, Little and Poggio, 1989) over patches of measured velocities to compute a smooth map of velocity over space. Some of these algorithms have also been implemented in neural networks (Wang, Mathur and Koch, 1989). These are attractive schemes since they employ techniques which can account for a number of perceptual phenomena, e.g. motion capture (Ramachandran and Anstis, 1983) and the aperture problem (Wallach, 1976, Marr and Ullman, 1981).

Unfortunately, none of these algorithms can deal with the recovery of SFM for transparent objects such as our rotating cylinder: vectors (with opposing direction) from the front and rear surface are assigned to one surface, and the averaging of velocities over a patch yields zero velocity. Evidently, an additional requirement for the successful application of these algorithms is the segregation of surfaces prior to smoothing. Recent work from our laboratory has demonstrated that many cells tuned for direction of motion in monkey striate cortex act as simple directional filters which are not influenced by the presence of dots moving in the non-preferred direction (Erickson, Snowden, Andersen, Treue, 1989). Thus as early as V1 two transparent surfaces moving in opposite directions will

excite different populations of neurons, thereby providing the first step for surface segregation based on direction of motion. Presumably, similar mechanisms allow for surface segregation based on speed as shown psychophysically in several studies (Ramachandran, Cobb and Rogers-Ramachandran, 1988, Andersen, 1989).

In conclusion, we have demonstrated the existence of a rather invariant point lifetime threshold for perception of 3-D structure-from-motion. Our data suggest that this threshold reflects the limits of accurate velocity perception and that accurate velocity measurements are critical for the computation of structure from motion. Furthermore, our results support the hypothesis that the visual system integrates information over space and time by computing a 3-D surface representation of the object. Such a process renders the visual system more flexible in its use of transient features for the SFM computation than current position-based schemes and is therefore more plausible. Finally, our data suggest that both velocity measurements and surface interpolation reflect global processes.

References:

Andersen G. J. (1989) Perception of three-dimensional structure from optical flow without locally smooth velocity. Journal of Experimental Psychology: Human Perception & Performance, 15, 363-371.

Borjesson E. and von Hofsten C. (1973) Visual perception of motion in depth: application of a vector model to three-dot motion patterns. Perception & Psychophysics, 13, 169-179.

Braunstein M. L. (1962) Depth perception in rotating dot patterns: Effects of numerosity and perspective. Journal of Experimental Psychology, 64, 415-420.

Braunstein M. L., Hoffman Donald D., Shapiro L. R., Andersen G. J. and Bennett B. M. (1987) Minimum points and views for the recovery of three-dimensional structure. Journal of Experimental Psychology: Human Perception & Performance, 13, 335 - 343.

Bülthoff H., Little J. and Poggio T. (1989) A parallel algorithm for real-time computation of optical flow. Nature, 337, 549-553.

Burr D. C. and Ross J. (1982) Contrast sensitivity at high velocities. Vision Research, 22, 479-484.

Clocksin W. F. (1980) Perception of surface slant and edge labels from optical flow: a computational approach. Perception, 9, 253-269.

Collett T. S. (1985) Extrapolating and interpolating surfaces in depth. Proceedings of the Royal Society, London, B 224, 43-56.

De Bruyn B. and Orban G. A. (1988) Human velocity and direction discrimination measured with random dot patterns. Vision Research, 28, 1323-1335.

Doner J., Lappin J. S. and Perfetto G. (1984) Detection of three-dimensional structure in moving optical patterns. Journal of Experimental Psychology: Human Perception & Performance, 10, 1-11.

Dosher B. A., Landy M. S. and Sperling G. (1989) Kinetic depth effect and optic flow. I. 3D shape from Fourier motion. Vision Research, 29, 1789-1813.

Downing C. and Movshon J. A. (1989) Spatial and temporal summation in the detection of motion in stochastic random dot displays. Investigative Ophthalmology and Visual Science (Supplement), 30, 72.

Erickson R. G., Snowden R. J., Andersen R. A. and Treue S. (1989) Directional neurons in awake rhesus monkeys: Implications for motion transparency. Society for Neuroscience Abstracts, 15, 323.

Green B. F. (1961) Figure coherence in the kinetic depth effect. Journal of Experimental Psychology, 62, 272-282.

Grzywacz N. M. and Hildreth E. C. (1987) Incremental rigidity scheme for recovering structure from motion: position-based versus velocity-based formulations. Journal of the Optical Society of America, 4, 503 - 518.

Grzywacz N. M., Hildreth E. C., Inada V. K. and Adelson E. H. (1988) The temporal integration of 3-D structure from motion: A computational and psychophysical study. In Organization of Neural Networks (Edited by von Seelen, W., Shaw, G., Leinhos. U. M.), 239 - 259.

Hildreth E. C. (1984) The computation of the velocity field. Proceedings of the Royal Society, London, Ser. B, 221, 189 - 220.

Hildreth E. C. and Koch C. (1987) The analysis of visual motion: From computational theory to neuronal mechanisms. Annual Review of Neuroscience, 10, 477-533.

Hoffman D. D. (1982) Inferring local surface orientation from motion fields. Journal of the Optical Society of America, 72, 888-892.

Hoffman D. D. and Bennett B. M. (1986) The computation of structure from fixed-axis motion: Rigid structures. Biological Cybernetics, 54, 71-83.

Horn B. K. P. and Schunk B. G. (1981) Determining optical flow. Artificial Intelligence, 17, 185 - 203.

Husain M., Treue S. and Andersen R. A. (1989) Surface interpolation in 3-D structure-from-motion perception. Neural Computation, 1, 324-333.

Johansson G. (1975) Visual motion perception. Scientific American, 232, 76-88.

Kelly D. H. (1979) Motion and vision. II. Stabilized spatio-temporal threshold surface. Journal of the Optical Society of America, 69, 1340-1349.

Koenderink J. J. and van Doorn A. J. (1986) Depth and shape from differential perspective in the presence of bending deformations. Journal of the Optical Society of America, A 3, 242-249.

Landy M. S., Dosher B. A., Sperling G. and Perkins M. E. (1988) The kinetic depth effect and optic flow. II. Fourier and non-Fourier motion. Mathematical Studies in Perception and Cognition, 88-4, NYU Report Series.

Lappin J. S. and Fuqua M. A. (1983) Accurate visual measurement of three-dimensional visual patterns. Science, 221, 480-482.

Lappin J. S., Doner J. F. and Kottas B. L. (1980) Minimal conditions for the visual detection of structure and motion in three dimensions. Science, 209, 717 - 719.

Longuet-Higgins H. C. and Prazdny K. (1980) The interpretation of a moving retinal image. Proceedings of the Royal Society, London, Ser. B, 208, 385-397.

Marr D. and Ullman S. (1981) Directional selectivity and its use in early visual processing. Proceedings of the Royal Society, London, Ser. B, 211, 151-180.

Mather G. (1989) Early motion processes and the kinetic depth effect. Quarterly Journal of Experimental Psychology, 41 A, 183-198.

McKee S. P. and Welch L. (1985) Sequential recruitment in the discrimination of velocity. Journal of the Optical Society of America, A 2, 243-251.

Miles W. R. (1931) Movement interpretations of the silhouette of a revolving fan. American Journal of Psychology, 43, 392-405.

Mitchison G. J. and McKee S. P. (1985) Interpolation in stereoscopic matching. Nature, 315, 402-404.

Morgan M. J. and Watt R. J. (1982) Mechanisms of interpolation in human spatial vision. Nature, 299, 553-555.

Nakayama K. (1985) Biological image motion processing: A review. Vision Research, 25, 625-660.

Nakayama K. and Silverman G. H. (1984) Temporal and spatial characteristics of the upper displacement limit for motion in random dots. Vision Research, 24, 293-299.

Newsome W. T. and Paré E. B. (1988) A selective impairment of motion perception following lesions of the middle temporal visual area (MT). Journal of Neuroscience, 8, 2201-2211.

Petersik T. J. (1987) Recovery of structure from motion: Implications for a performance theory based on the structure-from-motion theorem. Perception & Psychophysics, 42, 355-364.

Poggio T. and Koch C. (1985) Ill-posed problems in early vision: from computational theory to analog networks. Proceedings of the Royal Society, London, Ser. B, 226, 303-323.

Prazdny K. (1986) Three-dimensional structure from long-range apparent motion. Perception, 15, 619-625.

Ramachandran V. S. and Anstis S. M. (1983) Displacement thresholds for coherent apparent motion in random-dot patterns. Vision Research, 23, 1719-1724.

Ramachandran V. S., Cobb S. and Rogers-Ramachandran D. (1988) Perception of 3-D structure from motion: The role of velocity gradients and segmentation boundaries. Perception & Psychophysics, 44, 390-393.

Siegel. R. M. and Andersen R. A. (1988) Perception of three-dimensional structure from motion in monkey and man. Nature, 331, 259-261.

Siegel R. M. and Andersen R. A. (1986) Motion perceptual deficits following ibotenic acid lesions of the middle temporal area in the behaving rhesus monkey. Society for Neuroscience Abstracts, 12, 1183.

Snowden R. J. and Braddick O. J. (1989) The combination of motion signals over time. Vision Research, 29, 1621-1630.

Sperling G., Landy M., Dosher B. A. and Perkins M. E. (1989) The kinetic depth effect and identification of shape. Journal of Experimental Psychology: Human Perception & Performance, 15, 826-840.

Todd J. T., Akerstrom R. A., Reichel F. D. and Hayes W. (1988) Apparent rotation in three-dimensional space: Effects of temporal, spatial, and structural factors. Perception & Psychophysics, 43, 179 - 188.

Ullman S. (1984) Maximizing rigidity: The incremental recovery of 3-D structure from rigid and nonrigid motion. Perception, 13, 255-274.

Van Doorn A. J. and Koenderink J. J. (1982) Temporal properties of the visual detectability of moving spatial white noise. Experimental Brain Research, 45, 179-182.

van Doorn A. J. and Koenderink J. J. (1984) Spatiotemporal integration in the detection of coherent motion. Vision Research, 24, 47-53.

Wallach H. and O'Connell D. N. (1953) The kinetic depth effect. Journal of Experimental Psychology, 45, 205-217.

Wang H. T., Mathur B. and Koch C. (1989) Computing optical flow in the primate visual system. Neural Computation, 1, 92-103.

White B. W. and Mueser G. E. (1960) Accuracy in reconstructing the arrangements of elements generating kinetic depth displays. Journal of Experimental Psychology, 60, 1-11.

Williams D. and Phillips G. (1986) Structure from motion in a stochastic display. Journal of the Optical Society of America, A 3, P12.

Würger S. M. and Landy M. S. (1989) Depth interpolation with sparse disparity cues. Perception, 18, 39-54.

Yuille A. L. and Grzywacz N. M. (1988) A computational theory for the perception of coherent visual motion. Nature, 333, 71-74.

Zucker S. W. and Iverson L. (1986) From orientation selection to optical flow. Computer Vision, Graphics, & Image Processing.

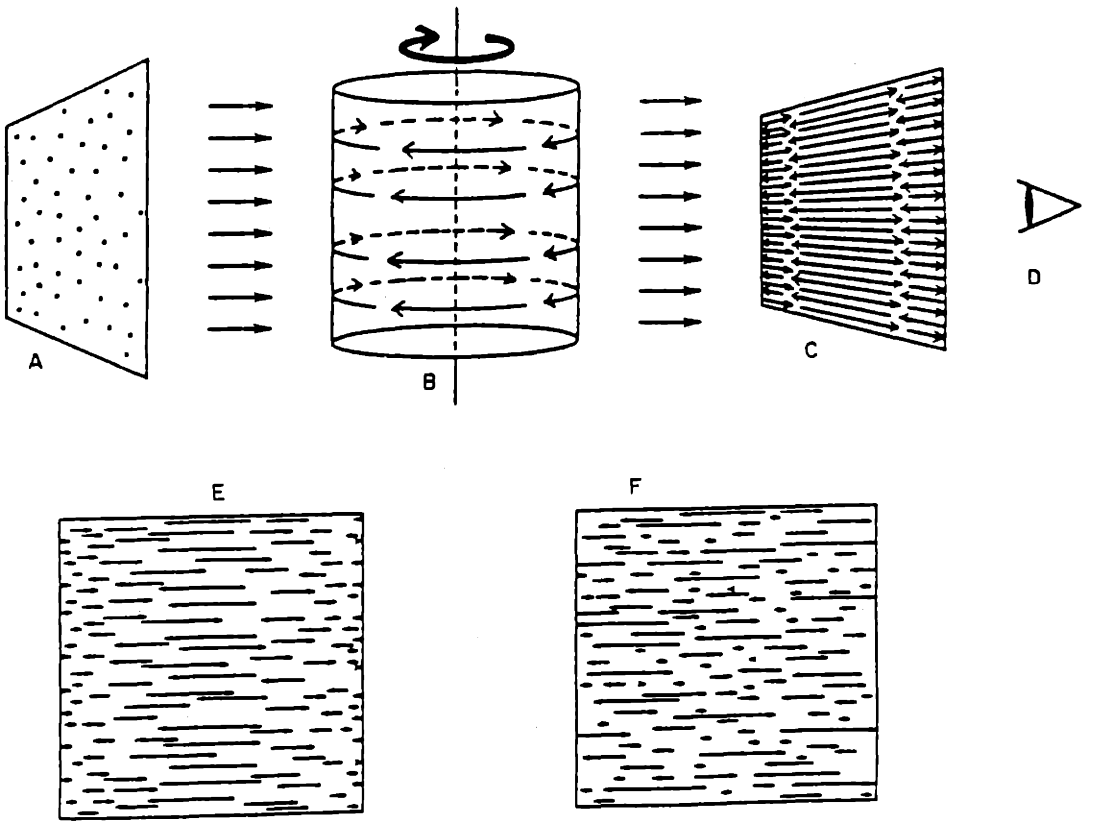


Figure 2.1

(A-D) Visualization of the algorithm used to generate the stimulus. Points are randomly plotted onto a square. They are then projected orthographically onto a transparent cylinder which is rotated. The orthographic projection of each point is then stored in the memory of a PDP 11-73 computer. Subjects viewed the movies on a Hewlett Packard 1311B CRT screen (phosphor P31). Viewing distance was 57 cm. The stimulus extended 6 x 6 angular degrees.
(E) Velocity field representing the structured stimulus. The velocity is small at the sides and high in the center, following a half cycle of a sinusoid along any horizontal line across the stimulus.
(F) Velocity field representing the unstructured stimulus. To generate this stimulus every vector from the structured stimulus was offset by a random amount within the stimulus boundaries.

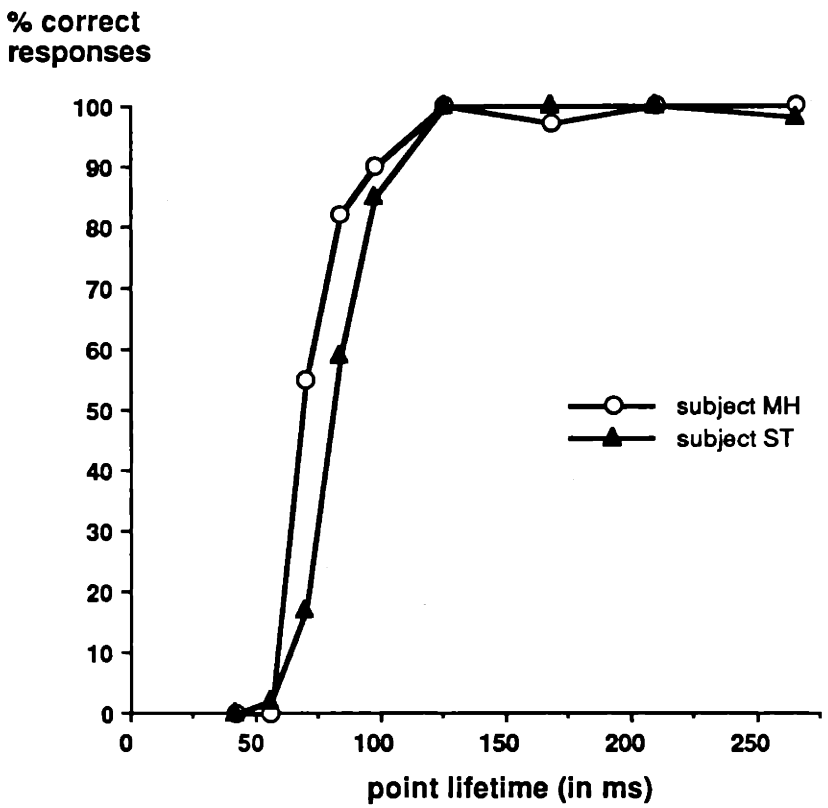


Figure 2.2

Percent correct performance in reaction time task plotted as function of point lifetime. Each data point represents between 38 and 42 trials. The curves follows a sigmoidal shape. The threshold (50 % of final height) are at ~ 69 / 81 ms. Because of the similarity between the two subjects all the following graphs show their averaged responses.

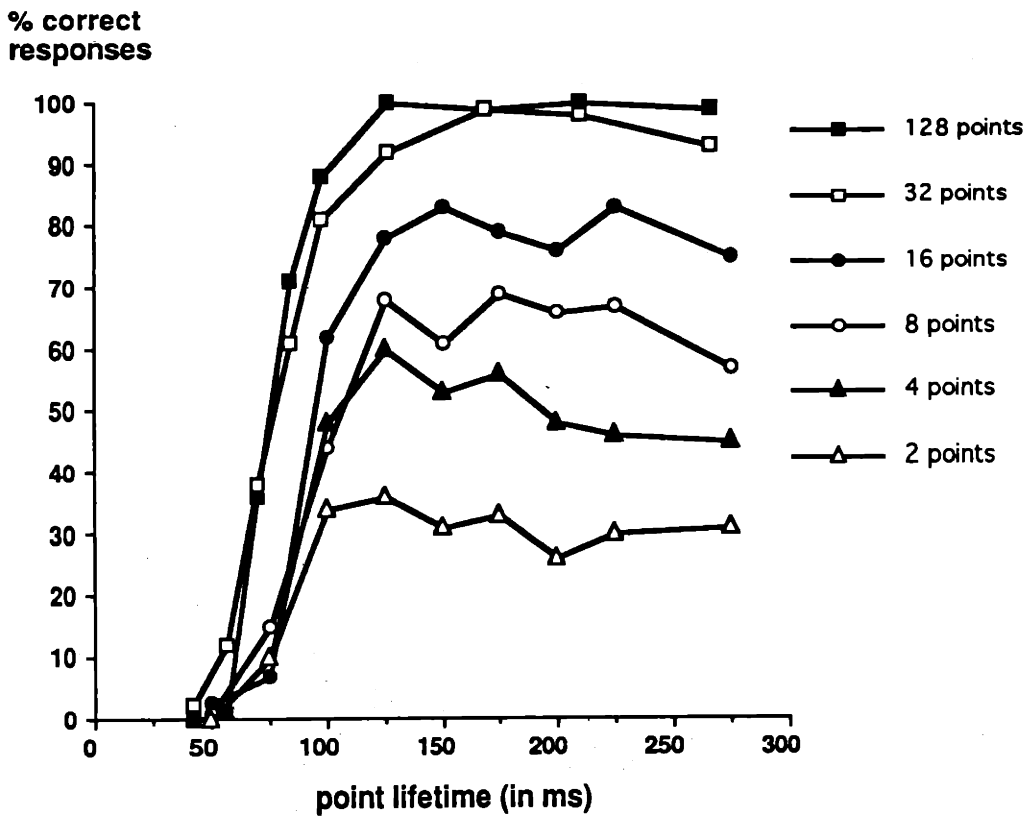
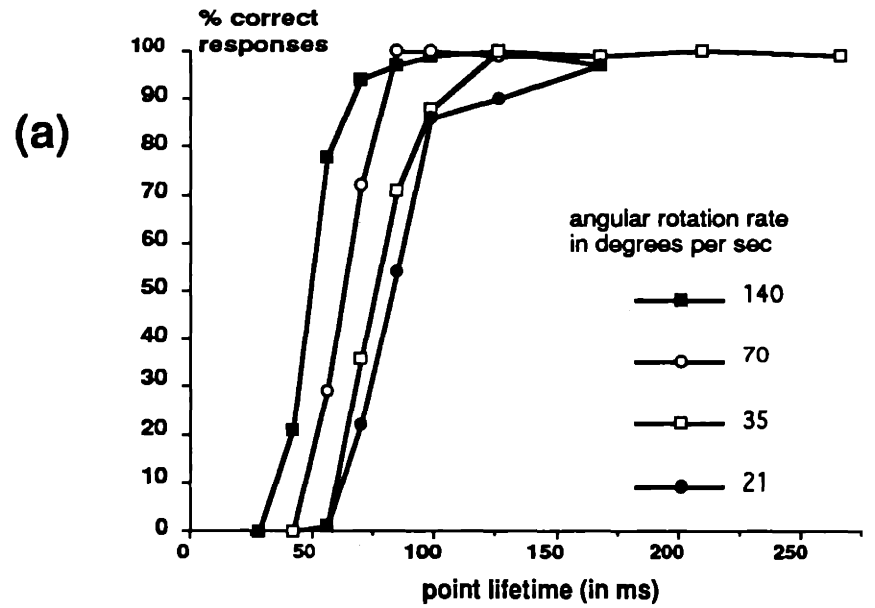


Figure 2.3

Percent correct performance in reaction time task at different numbers of points. Note the similarity in thresholds compared to the wide variation in the heights of the plateaus.



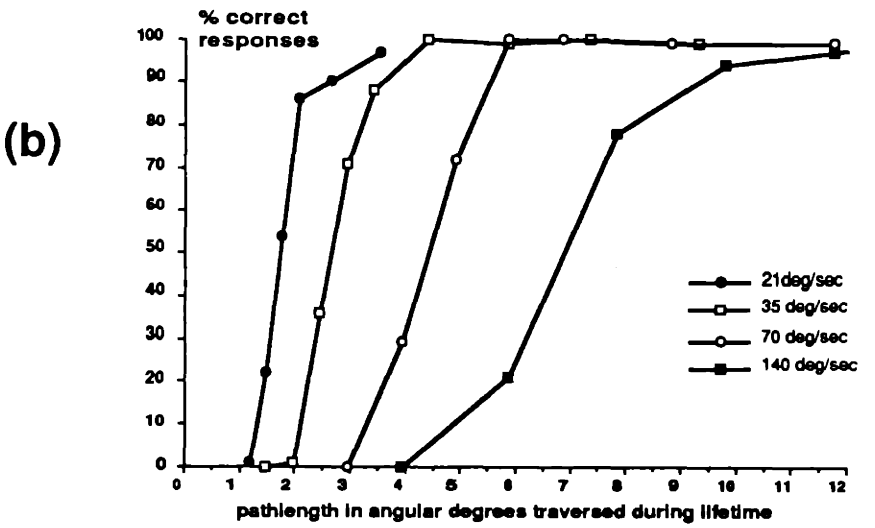


Figure 2.4 a+b

for figure legends see next page

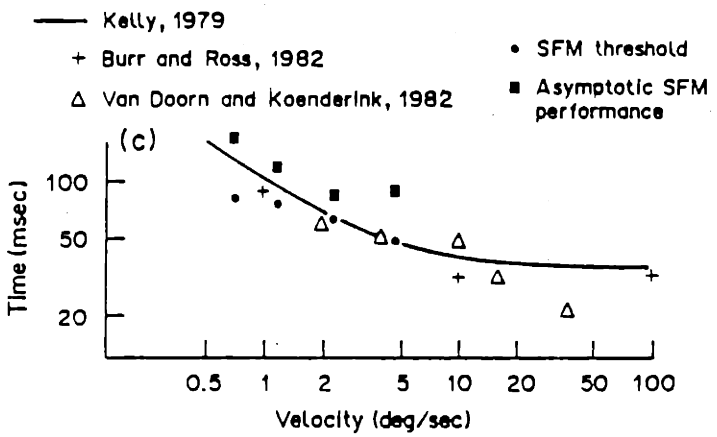


Figure 2.4 c

(a) Percent correct response in reaction time task plotted as function of rotation rates. Note the similarity in curve shapes.
(b) Data from Fig. 4a replotted as function of pathlength in angular degrees travelled by individual points. Note the large shift in thresholds as compared to Fig. 4a.
(c) Threshold for SFM perception (filled circles) and the lifetimes at which SFM performance peaks (filled squares) as a function of 2-D velocity in the display plotted together with optimal temporal intervals for the inputs to hypothesized direction selective subunits as estimated by Nakayama, 1985 (his Figure 6).
The solid line comes from calculating the optimal temporal intervals from the peak spatial and temporal frequency contrast sensitivity for detection (from the sine wave data of Kelly, 1979). Crosses come from data measuring the contrast sensitivity for direction discrimination (Burr and Ross, 1981) analysed in the same way as Kelly's data. The triangles are derived from experiments requiring the observer to see coherent motion of random dots in a field of dynamic visual noise (Van Doorn and Koenderink, 1982).
It should be noted that despite wide differences in the experimental paradigm and observer task, estimates of optimal timing for a given velocity show considerable similarity.

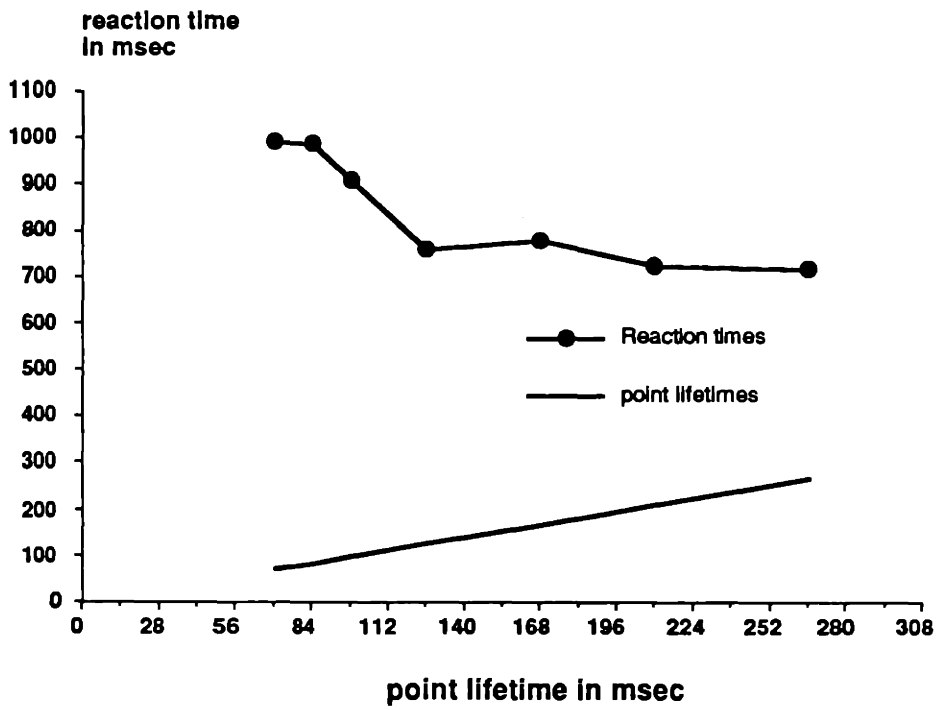


Figure 2.5

Reaction time as a function of point lifetime. Point lifetime is also plotted to allow easy comparison with reaction time. Note that the reaction time is always many times longer than the point lifetime.

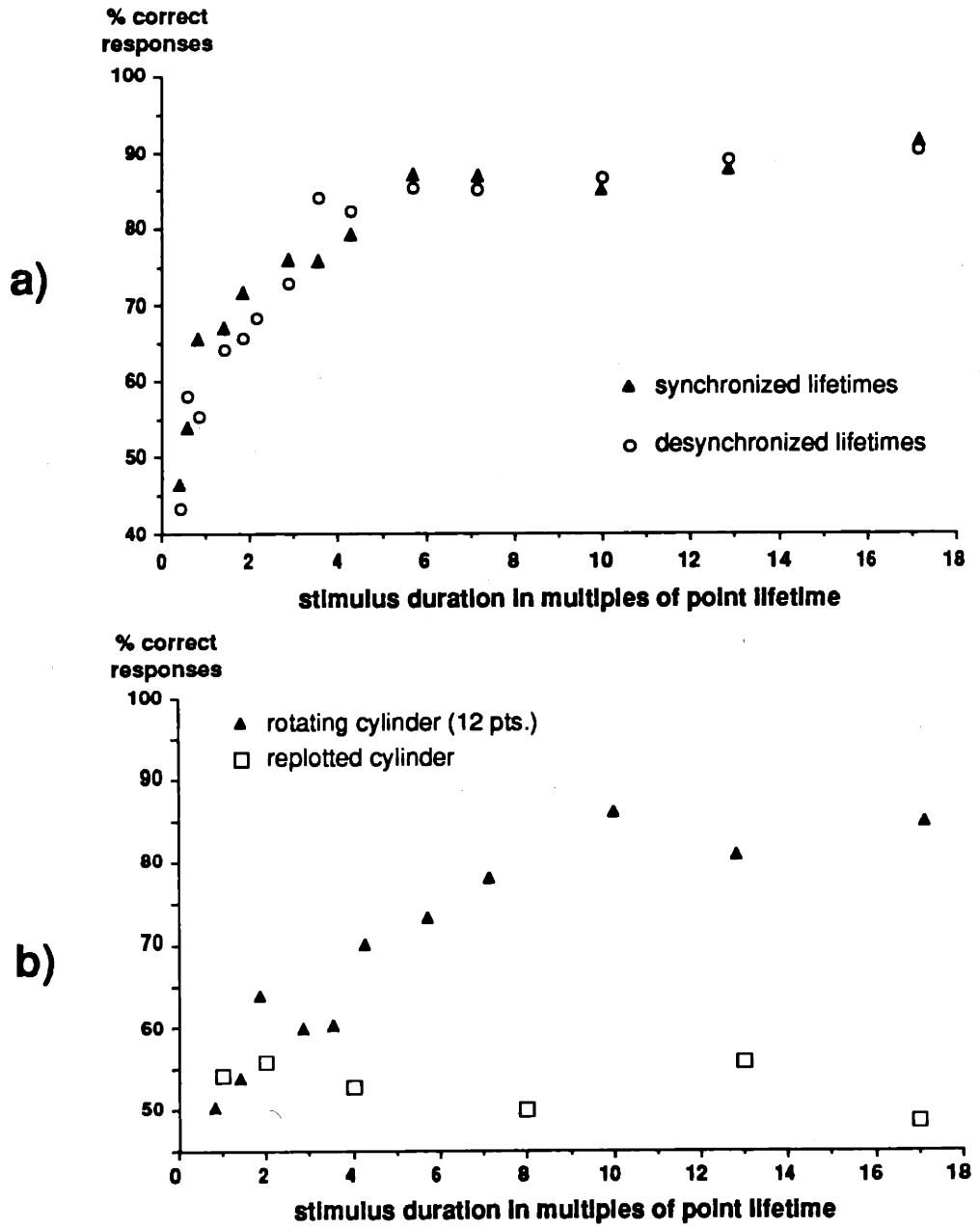


Figure 2.6

(a) Percent correct responses in 2AFC task plotted as a function of stimulus duration. Note the long build up of performance.

(b) (open symbols) Percent correct responses using 12 points.
(filled symbols) Percent correct when the points were repeatedly travelling along the same path.

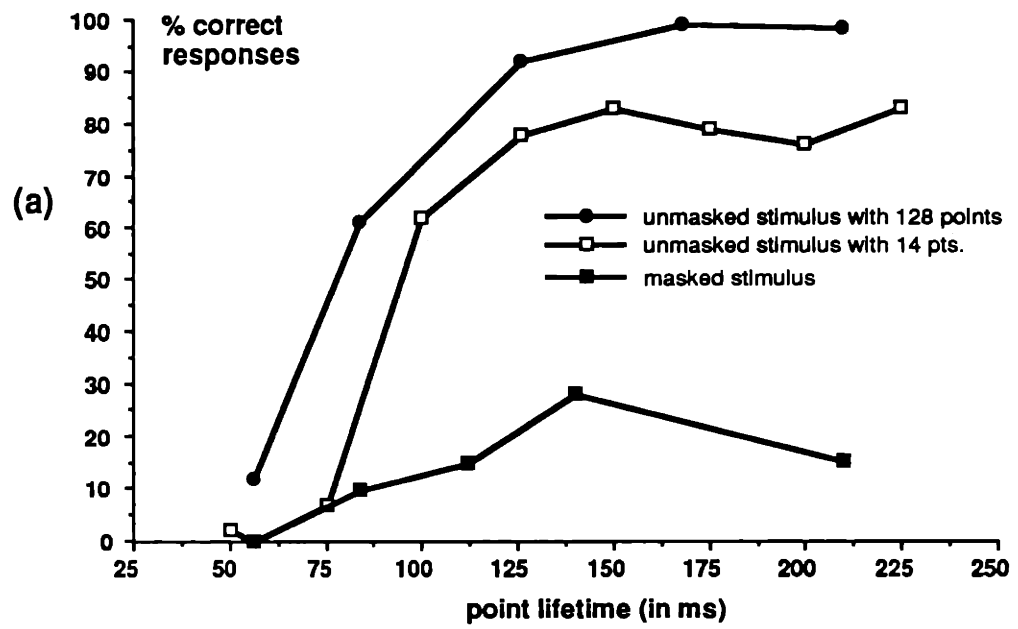
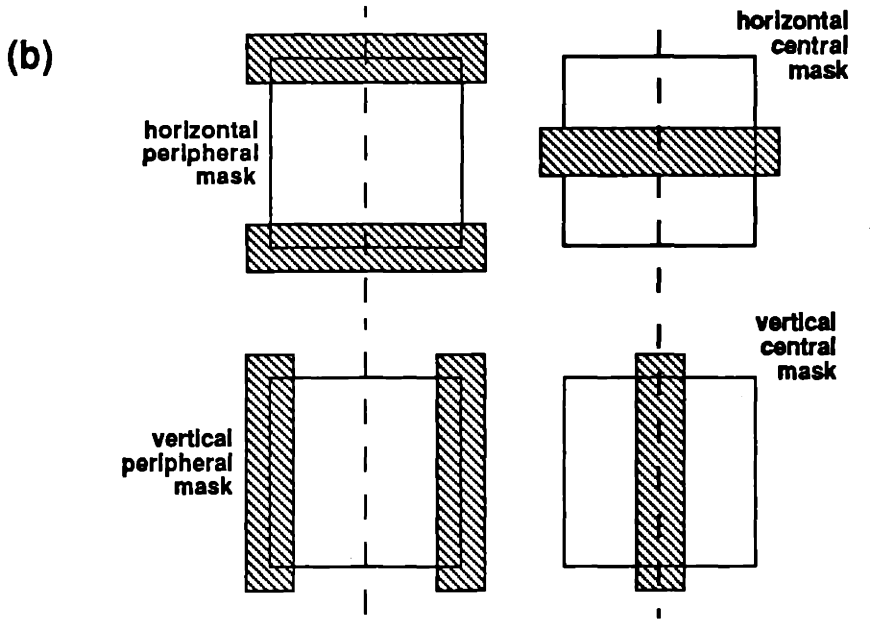


Figure 2.7 a

Performance in reaction time task comparing unmasked stimuli with masked stimuli. The number of points in the unmasked part of the stimulus averaged 32 to allow easy comparison with our data from the unmasked stimulus.
(a) Performance when only ~ 11 % of the stimulus was visible (average of ~ 14 points) compared to performance using full size stimulus with either same number of points (open squares) or with the same dot density (filled circles).
(b) Shape and location of masks used to cover 25 % of the stimulus. The two horizontal masks cover redundant areas of the stimulus while the vertical mask cover non-redundant areas.
(c) Comparison of the effect of a 25 % mask of different shape and orientation.
Open circles: Performance using central vertical mask
Open squares: Performance using central horizontal mask
Filled circles: Performance using peripheral vertical mask
Filled circles: Performance using peripheral vertical mask.
Dotted line: Performance with the unmasked cylinder using 32 points (from Fig. 3).
Note the difference in performance between stimuli in which a non-redundant part was masked away (circles) and when a redundant part of the stimulus was invisible (squares).



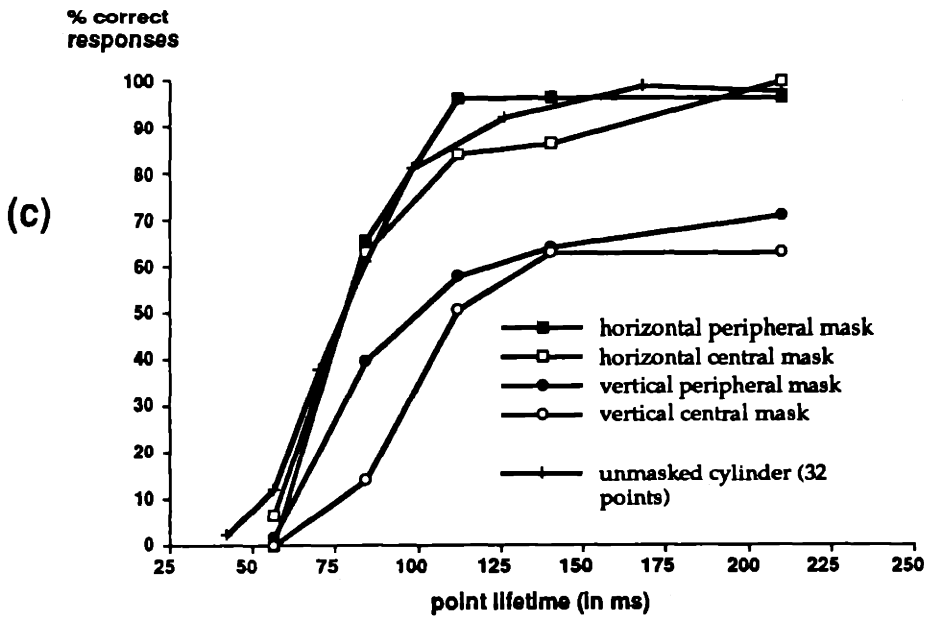


Figure 2.7 b + c

(legend on previous page)

Chapter 3

Structure from Motion : Perceptual Evidence for Surface Interpolation

Introduction

When the visual system is trying to recover the three-dimensional (3-D) shape of an object it is presented with the problem of only having access to the two-dimensional (2-D) images projected onto its retinae. Besides the disparity between the images in the two eyes a multitude of monocular cues (like shading, texture, occlusions, and motion) enable the visual system to perform the task despite this limitation. Extracting the three-dimensional shape of objects from the relative motion of their parts is called structure-from-motion (SFM) or the kinetic depth effect and has been studied extensively since its first description by Miles (1931).

Computational studies on extracting SFM suggest two possible approaches. The visual system could use the change of the relative 2-D positions of object features (like edges, corners, line terminations or texture elements) over time in a position-based approach. In a velocity based approach on the other hand it would analyze the pattern of 2-D velocities generated by the stimulus features. Our companion paper provides a review of the various computational implementations of these two approaches (Hildreth, Ando, Andersen, & Treue, 1992).

In a previous study (Husain, Treue & Andersen, 1989; Treue, Husain & Andersen, 1991) we have presented psychophysical evidence for the use of velocity in the extraction of SFM in humans (see also Dosher, Landy, & Sperling, 1989). We have further argued that the visual system is able to integrate the depth information derived from the individual stimulus features spatially through a process of surface interpolation and that such an interpolation mechanism facilitates the integration of motion information over time.

We use the term interpolation informally to describe a process in which the object depth values extracted at the stimulus features are used to reconstruct a complete surface in depth which fills-in depth values between the known depth values. The companion paper provides a discussion of various possible implementations of such a process.

SFM without a surface interpolation process would use the motion of stimulus features to extract the corresponding depth values only, leading to a wire-frame representation of the object that makes no implicit assumptions about depth values between features.

In the experiments presented here we investigate the proposed surface interpolation process psychophysically and test several predictions which can be derived from our surface interpolation hypothesis. We further show that a number of previous observations regarding the extraction of SFM can be accounted for by our surface interpolation hypothesis. In the companion paper (Hildreth, Ando, Andersen, & Treue, 1991) we demonstrate how to incorporate both surface interpolation and temporal integration into computational algorithms for the recovery of 3-D SFM.

The most prominent characteristic of any spatial interpolation process is that it "fills-in" featureless areas. In the first experiment we test the ability of human observers to detect such featureless stimulus areas as a function of their size. We show that as predicted by a surface interpolation process, subjects have great difficulties in detecting the presence of these featureless areas on the object.

One of the advantages of using a surface interpolation process

in the extraction of SFM is that the extracted surface incorporates the information derived from individual features and preserves that information even after the disappearance of any individual feature. This allows features that appear separated in time to contribute to the extraction of the same object. The wire-frame approach described above only represents the depth values of the currently present features and thus shows no such temporal integration. The prime purpose of the extracted surface could thus be to link features over time without actually being used in the neural representation of the object. Alternately, if the information from individual stimulus features is only used to allow for the extraction of the interpolated surface but is not preserved explicitly beyond the interpolation stage then the internal representation of the observed object is a surface rather than a collection of individual elements. By putting the behavior of the individual features and that of the object's surface in disagreement we are able to demonstrate in the second experiment that the 3-D percept indeed follows the behavior of the surface rather than that of the individual features, in agreement with a rather fundamental role for the interpolated surface.

Given the richness of the visual world outside the psychophysical display a surface interpolation process used to recover object shapes should incorporate knowledge about object boundaries to spatially limit the interpolation process. In a final series of experiments we relate various observations by Ramachandran et al. to a SFM process that uses surface interpolation and boundary constraints.

Finally we argue that recent findings concerning the processing of motion transparency in the visual cortex of awake behaving monkeys puts the process of segmenting the front and

back surface of transparent objects (like those used in our experiments) as early as Area V1 of primate visual cortex.

Methods

General Aspects

The basic stimulus used in all of the experiments described here is a moving random dot pattern representing the parallel projection of a rotating transparent cylinder covered with points (Fig. 3.1A, B & C). This stimulus is generated on a PDP 11-73 computer and then presented to the subjects (Fig. 3.1D) on a CRT screen in a dimly lit room. The subjects sit without restraint and view the screen binocularly from a distance of 57cm. All experiments described here used a two-alternative forced choice procedure and the subjects held a box with two buttons to record their responses.

An important characteristic of our stimulus is the use of limited point lifetimes. All the dots are present at specific positions on the cylinder only for a predetermined number of frames and are then randomly repositioned. Thus the projected image consists of individual points moving only through short trajectories. This allows us to limit the amount of information an individual feature can contribute to the recovery of the perceived 3-D shape of the object. It also enables us to keep the 2-D density distribution constant across the display throughout the rotation of the cylinder (for a more in-depth discussion of our stimulus and its generation see Treue, Husain & Andersen, 1991).

The parallel projection of a rotating cylinder generates a 2-D flow field with a velocity distribution described by a half sinusoid (Fig. 3.1E) in which stimulus features in the middle of the display move at a high velocity that drops to 0 at the edges of the display where the dots reverse their direction and move along the opposite surface.

Since it is relevant especially for our experiment 2 it should be pointed out that the perceived direction of rotation for parallel projected objects is ambiguous and can reverse spontaneously during observation.

Experiments

Experiment 1:

Using a surface interpolation mechanism would not only allow the visual system to integrate information from points being present in the display at different times but would also allow the visual system to reconstruct complete objects, i.e. surfaces from sparse data. As mentioned in the introduction one would expect a surface interpolation mechanism to fill in data between the points in our stimulus. Such a filling-in process should lower the visual system's ability to detect areas without features. We set out to demonstrate this side-effect of interpolation by measuring subjects' inability to detect the presence of masked parts or 'holes' in the surface of a rotating cylinder as a function of the size of the occluded area.

Methods

Subjects were presented with either a complete rotating cylinder or a rotating cylinder with a cut-out part. We investigated two mask conditions: In the first we masked one of four possible locations on either of the two surfaces of a rotating cylinder. These masked areas were stationary and centered on the middle of the four quadrants of the stimulus as sketched in Fig. 3.2a. In the second condition the masked area was stationary on the surface of the cylinder and thus moved across the display during the stimulus presentation. The mask was

randomly placed within a part of the cylinder that would not rotate from the front to the back or vice versa during the stimulus duration.

In both conditions the mask covered only one of the two surfaces, because otherwise the detection of a mask would amount to nothing more than detecting stimulus areas void of points. For the same reason, the distribution of points was varied so that the surface patch opposite to the mask contained twice as many points as usual, thus guaranteeing an even dot density across the stimulus. This density control could be easily implemented because of our use of limited lifetimes. The projected size of the hole present in the second mask condition was kept constant while rotating around the cylinder so as to allow a better comparison to the first mask condition.

A further control was necessary in the first mask condition. Points on the surface of the cylinder disappeared when they rotated into/behind the stationary mask in this condition and reappeared on the opposite side of the mask. This could serve as a cue to the subjects about the presence and location of the mask. To invalidate this spurious cue we positioned 'virtual' masks in all four quadrants of the masked as well as of the unmasked cylinders. These virtual masks behaved like real masks in that points crossing their boundaries were replaced to the opposite side of these masks but they differed from real masks in that they contained moving dots rather than masking them. Through this manipulation virtual lines generated by disappearing and reappearing dots were present in all stimuli and could not serve as cues to the presence of a featureless area.

Subjects were presented with the stimuli which lasted 2 secs and rotated at an angular rotation rate of 50°/sec. The

stimulus was made up of 125 points on each surface. The points lifetime was 200 msecs, long enough for a strong impression of SFM (see Treue et al. 1991) and short enough to introduce point disappearances and appearances that masked the edges of both the real mask as well as the 'virtual' masks. These mask were squares of 2.25 to 20.25 angular deg^2 (the stimuli subtended 10x10 angular degrees). The number of points masked thus ranged from about 3 to over 25 points. Subjects were instructed to press one of two buttons after the end of the stimulus to indicate whether they perceived the stimulus as a complete rotating cylinder or if they detected a mask/hole.

After a block of trials that typically consisted of 80 stimulus presentations, subjects were also asked which direction they perceived the cylinder as rotating[1]. With this information it was possible to plot performance separately for holes in the perceived front or perceived back of the cylinder.

Figure 3.3 shows the results of the two experiments we performed. Figure 3.3A plots subjects' performance when presented with cylinders containing holes that moved with the rotation. Figure 3.3B plots the results of using masks that were stationary in 2-D. In both cases hole width in angular degrees is plotted on the x-axis while the two curves represent the performance for the hole being either on the perceived front or perceived back of the cylinder respectively. Chance in our two-alternative-forced-choice paradigm was 50% and is denoted by a horizontal line.

Two results are apparent from the data. To be able to detect the presence of a hole or mask at the 83% level the size of the

[1] Fortunately subjects generally have such a strong bias for one direction of rotation that during our short stimulus durations they never perceived reversals of the perceived direction of rotation and also generally did not perceive different directions of rotations for the 80 trials in a block.

mask has to be about 10 deg^2, i.e. it has to cover nearly half of the stimulus quadrant (5*5 deg) it is placed in^2. This indicates that although subjects are able to easily segment the front and back in our transparent stimulus they have great difficulty evaluating the completeness of each surface. This result supports our hypothesis, that the final percept of 3-D shape is more closely related to an interpolated surface representation than to the 3-D structure of individual features (see experiment 2).

Furthermore for the stationary mask there was a curious dependence of performance on the surface the mask was perceived to cover. Subjects were very poor at detecting even very large masks when they covered the perceived back surface of the cylinder. This difference between the two surfaces was not present for the hole that moved with the cylinder. Perceptually this seems to reflect the fact that a stationary "hole" in the back of the cylinder can only be achieved by the presence of a stationary non-transparent object within the cylinder. Given that subjects do not see such an occluder in our experiment they tend to ignore the possibility of a stationary mask covering the back surface.

Experiment 2:

After strengthening the argument for the presence for surface interpolation in SFM we address the question of how fundamental a role this surface interpolation plays in the mental representation of the extracted object. Two possibilities come to mind. The surface interpolation might simply be used to allow the visual system to recover the 3-D positions of individual stimulus elements even when they are temporally separated

2 In Figure 2b+c we sketched the effect of a mask of that size on the 125 points present in one surface of the cylinder.

without playing a role in the internal representation of the observed object. In such a scheme the final representation of the stimulus would be as a group of points or a wire-frame style object positioned in 3-D. On the other hand, it is possible that the information from the individual features is only used to interpolate the surface, and that the object is ultimately represented through its surface rather than as a collection of individual elements in space or as a wire-frame style representation. If such a scheme were indeed employed by the visual system then the extraction of SFM should be determined by the behavior of the object's surface rather than by the behavior of the individual points.

Our stimulus allows us to perform an interesting variation to distinguish between these two schemes. Because we use a rotationally symmetric object in orthographic (parallel) projection the assignment of the front and back surface is arbitrary and in fact sometimes reverses spontaneously during viewing (similar to the Necker Cube). Unlike the Necker Cube this switch is not accompanied by a change in the object's surface shape or position, but rather only by a change in direction of apparent rotation. Thus the perceptual reversal of a Necker Cube represents a physically unlikely event while the perceptual reversal of our rotating cylinder corresponds to a change in direction of rotation. We were thus interested in comparing these apparent changes in rotation direction with the perception of "real" (physical) changes in direction of rotation.

We presented subjects with a rotating cylinder in which all points reversed their direction of motion in synchrony after randomly chosen periods of time. The subjects' task was to indicate every perceived reversal of rotation direction by pressing a button. As a control we also measured the number of

perceived reversals in a cylinder in which the individual points never reversed their direction of motion.

The results are plotted as the average of 5 subjects in Figure 3.4. The left bar indicates how often the points in the display changed their direction of motion during the stimulus duration. The middle bar is the number of reversals of direction of rotation perceived by the subjects. The right bar represents the control measure using a smoothly rotating cylinder. Two findings are immediately obvious. First the number of perceptual reversals in the smoothly rotating cylinder is rather low (in agreement with what the subjects reported in experiment 2). Secondly the subjects only saw about half of the "true" reversals as such.

Two questions arise from this finding. How could the subjects miss so many of the reversals and how did they perceptually interpret the reversals which they did not see as such ? When debriefing the subjects after the experiment all reported two different percepts. Sometimes they saw the rotating cylinder reverse its direction (and pressed the button as instructed). At other times the cylinder seemed to stop momentarily and then continue to rotate in the same direction as before the stop.

To account for this percept one has to assume that the visual system changed the assignment of all points from their current surface (front or back) to the opposite surface whenever they reversed their direction of motion. Only then can the cylinder as a whole be perceived to rotate in the same direction after the direction of motion of all the individual points is reversed. Thus the answer to the two questions posed above is: The subjects did see all instances of reversals of direction of motion of the individual points, but interpreted only some of them as a

reversal of the rotation of the overall cylinder. Such a percept seems only possible if the object is internally represented as a 3-D surface rather than as a group of dots in space, because none of the subjects reported seeing the individual dots move through space from one surface to the other when the percept of "stopped" motion occurred. This is quite different from the perceptual reversals of the Necker Cube which is clearly interpreted as a change in 3-D location of the stimulus features.

It could be argued that the case presented to the subjects above is special in that a wire-frame representation of the cylinder could be maintained through the reversal of direction of motion since the 3-D distances between all points remain the same and all that happens is a depth reversal. One can think of the cylinder before and after the percept of stopped motion occurred as mirror images of each other. To control for that possibility we introduced a further modification to our stimulus.

We desynchronized the reversals of the individual points in time, i.e. between any two frames of our stimulus presentation only a small proportion of points reversed their direction of motion. If the visual system represents the points in the stimulus at their individually computed locations in space the stimulus would be interpreted as an entirely non-rigid cloud of dots since the two dots in any pair will sometimes move in the same and sometimes in the opposite direction. A wire frame representation of the stimulus would not just switch between two depth-reversed states but rather would be constantly changing since individual points would jump from the front to the back and vice versa in the process changing their 3-D distance to all other elements of the object.

When we presented this stimulus to naive observers they all

reported the percept of a smoothly rotating hollow cylinder. In fact we never found it possible to perceive anything but a smoothly rotating cylinder. The only way to tell this display from our regular cylinder as used for the control above (right column in Fig. 3.4) is to carefully track an individual point throughout the stimulus duration.

The percept of a smoothly rotating rigid cylinder in the face of the highly non-rigid physical stimulus can only be accounted for if the object is represented as a surface without an explicit representation of the individual features that contributed to it.

Various Observations and Demonstrations:

In our previous studies (Husain et al. 1989, Treue et al. 1991) we investigated how precisely, how fast, and under which conditions humans can distinguish between a structured stimulus (the parallel projection of a rotating cylinder) and a control stimulus (generated by randomly shuffling the motion vectors of the structured display). But human perception of SFM is not always veridical and a variety of perceptual demonstrations (besides the one we documented above) have been documented in which subjects reported the percept of rotating cylinders or other objects when the display in fact was not physically consistent with such a percept.

Below we will describe several of these demonstrations as well as variations on them we developed and discuss them in light of our proposed surface interpolation process.

Segmenting multiple-surface displays

Any surface interpolation used to recover two surfaces present at the same image location (as is the case with our transparent cylinder) has to first segment the surface features

based on the surface on which they lie before performing an interpolation. If a single surface would be interpolated simultaneously through points laying on both the front and back of the cylinder the result would be either a flat surface or a highly convoluted, permanently changing shape, depending on the amount of smoothing performed. For rotating objects like the cylinder used in our experiments such a segmentation process would be relatively easy since the two surfaces move in opposite directions. Recordings from Area V1 of the awake behaving monkey suggest a physiologically plausible implementation of such a segmentation process. Snowden, Treue, Erickson, & Andersen (1991) have recently demonstrated that transparent moving random dot patterns activate two separate populations of direction tuned cells, one for each of the two opposite directions present in the display. On average, the response of the cells tuned to one direction is affected very little by the presence of points moving in the opposite direction. It thus seems that the visual system segments the two surfaces as early as V1 and this information might be used to guide the surface interpolation process. This line of reasoning is further supported by the results of our experiment 2. Given the "blindness" of direction-selective V1 cells for the presence or absence of their anti-preferred direction they would not be able to distinguish between our smoothly rotating cylinder and the two stimulus variations employed in experiment 2. These cells would only "see" one surface of the cylinder and points that belong to that surface and then reverse their direction of motion would simply "disappear". Such a point will then enter the group of points stimulating cells tuned for the opposite direction.

A visual illusion reported by Ramachandran, Cobb and Rogers-Ramachandran (1988) might be interpreted as evidence

for a segmentation process more powerful than the one based just on opposite directions as described above. These researchers generated two displays representing two coaxial transparent rotating cylinders. In one display one cylinder had a smaller radius and its rotation speed was increased so that the 2-D velocity in the middle of the display was the same for both cylinders. Observers report a percept in which the two cylinders' surfaces seems to coincide in the middle and are separated in depth towards the edges of the display. In a variation of this display the two cylinders were of the same radius but one was rotated at twice the rotation rate as the other. Rather than perceiving the two cylinders as coinciding in depth subjects reported a separation of their surfaces such that the faster rotating cylinder seemed to bulge out more in depth (see the companion paper (Hildreth et al 1991) for a more detailed description, analysis and figures). Since Ramachandran uses infinite point lifetimes and distributes points randomly on the cylinder (i.e. in 3D) rather than on the display these demonstrations contain an uneven density of points as a confounding depth cue. We were able to replicate the results even after removing density gradients in the display by plotting the points randomly in 2D and by the use of limited point lifetimes. These illusions might be interpreted as suggesting that the visual system can segment surfaces not only by opposite directions but also by their local 2-D speed[3]. Such an inference might be premature though because we demonstrate in the companion paper (Hildreth et al., 1991) that an algorithm that uses relative motion of features to determine their relative depth will generate comparable results even without surface interpolation and thus without any segmentation at all. The perceived segmentation of surfaces in these demonstrations might therefore not be due to an early segmentation process

[3] I think Ramachandran actually suggest something like that.

based on speed differences but might simply reflect the workings of a SFM process that assigns depth based on relative speed between stimulus features.

In relation to the multi-surface percept evoked by these stimuli it should be noted that subjects report difficulties perceiving all 4 surfaces (front and back of the two coaxial cylinders) at the same time. Rather they segment the attended surface (front or back) while perceiving a much less clear segmentation of the respective other surface. This is similar to findings by Andersen (1989) who reports that subjects can only detect up to three superimposed transparent surfaces moving in depth at a time. Since subjects' fixation was not monitored in any of these studies it is possible that maximally three surfaces were perceived since visual tracking could result in the image of one random dot pattern being stationary on the retina and two patterns moving in opposite directions (thus again opening the possibility that the segmentation even in this case is based on opposite directions of motion).

Effects of boundaries on SFM interpretation

Several studies have described displays in which boundaries influenced the 3-D interpretation of moving random dot patterns.

Ramachandran (1988) uses two superimposed random dot patterns moving with constant speed in opposite directions. When points reach the edge of the display they reverse their direction (they "bounce" off the edges). Ramachandran reports that subjects perceive the display as a rotating cylinder, rather than two flat planes. In a related demonstration Ramachandran masks the projection of a rotating cylinder so that only a triangular or rectangular section is visible. He reports that

subjects describe a percept of a complete rotating cone or cylinder, respectively, rather than that of a masked, incomplete cylinder.

Thompson (1991) also reports that under certain conditions a rectangular patch of random dots moving in one direction at constant speed surrounded by random dot patterns moving at a different direction or at different speeds can result in the percept of a rotating cylinder. Royden (1988) reports a similar finding for random dot patterns moving within a rectangular patch surrounded by a stationary random dot field.

Nakayama has recently suggested a framework which could help in interpreting these perceptual demonstrations. To capture the relationship between a perceived boundary and its two abudding surfaces he introduced the terms intrinsic and extrinsic. A boundary is intrinsic to a given surface if it is physically connected to the surface. Since the boundaries in all the demonstrations described above which generated percepts of rotating cylinders were perceptually intrinsic to the cylinders we set out to generate a display which included extrinsic boundaries.

The display is based on Ramachandran's observation that masking the sides of a vertically rotating cylinder results in the percept of a cylinder of smaller diameter, i.e. higher curvature. Our display contained four moving random dot patterns. Figure 3.5 is a single frame out of the sequence of frames displayed on a computer screen. Figure 3.5A is the parallel projection of a transparent rotating cylinder very similar to the display we used for experiments 1 and 2. Figure 3.5C is the same cylinder partially covered by a dark mask. Figure 3.5D represents a display similar to the one in 3.5C except that the mask is invisible. Figure 3.5B finally represents the parallel projection of

a cylinder of a width equal to the width of the random dot patterns in 3.5C and D.

When subjects describe their percepts of display C and D they report that C seems to be part of a masked cylinder similar to the one in A. Random dot pattern D, although physically identical to pattern C, is perceived as a cylinder of smaller diameter more like B (although generally not as highly curved as in B).

Just as described by Nakayama and colleagues (Nakayama, Shimojo, and Silverman, 1989; Shimojo, Silverman, and Nakayama, 1989; Nakayama, and Shimojo, 1990) for the occlusion of surfaces perceived in depth the visual system seems to distinguish between intrinsic and extrinsic surfaces in the recovery of SFM. The companion paper (Hildreth 1991) as well as Ando (1992) describes ways in which boundaries can interact with a surface interpolation mechanism to account for our observations.

General discussion

In the experiments presented here we strengthened the case for the involvement of a process of surface interpolation in the recovery of 3D SFM in the human visual system. Our findings suggest that the role of such a process goes beyond being simply a means for recovering the depth of individual feature elements presented in temporal separation. Rather the internal representation of the object in the visual system seems to be as an object described by its surface and not as a cloud of individual features.

Such a representation would provide an easy way for integrating other cues for depth perception which are often surface-based, like shape-from-shading, texture gradients and even binocular disparity, since there have been several reports for surface interpolation in stereoscopic depth perception (Mitchison & McKee, 1985; Mitchison & McKee, 1987; Würger & Landy, 1989).

This leaves the question of how we perceive objects that rather than having a distinct surface represent a volume of points. Two explanations seem possible. Volumes could be represented in an onion skin like fashion as layers of surfaces. Alternatively, in the absence of distinct surfaces the visual system could resort to tracking individual groups of points to determine the axis of rotation as well as the range of 3D distances from this axis present in the image. The possibility that the visual system could switch to another, non surface-based approach to extracting depth from motion seems not unlikely given that no surfaces can be recovered in instances like Johanson's biological motion displays (Johanson, 1975). In these cases heavy constraints derived from our knowledge of

how humans move allow a very accurate representation of the 3D motion from just a few strategically located moving points.

In the absence of such constraints Todd et al. (Bressan and Todd, 1990; Norman and Todd, 1992) using wire-frame objects in rotation find poor performance and little evidence for temporal integration. They show that subjects perform poorly on non-surface based tasks, as for example the estimation of 3-D line length while they show good performance using similar stimuli when comparing the slant of surfaces formed by two intersecting lines in depth.

Physiological Implementation

An important issue not addressed directly in the experiments reported here is the question of what kind of information the visual system uses as input to the surface interpolation process. This is relevant in light of current computational approaches that use either the changes in the relative 2-D positions of object features (position-based approaches) or the velocity field of the projected object (velocity-based approaches). Previously we have provided strong evidence that the visual system uses a velocity-based scheme (Treue et al., 1991). Thus the algorithm presented in the companion paper uses such an approach. The specific implementation presented tracks the velocity of individual features. This is a computational convenience. It should be pointed out, however, that it is not easy to translate such a scheme directly into biological hardware since the visual system with its stationary receptive field cannot really track individual features in dense random dot patterns directly[4]. Rather the direction-selective neurons in the visual cortex act as spatio-

[4] In fact that is the original motivation for Nakayama & Tyler's (1981) introduction of moving random dot patterns into studies of the visual system.

temporal filters that generate a representation of the optical flow present in the retinal images. Through the aperture provided by the individual receptive fields the visual system is already performing a smoothing operation on the visual input since an individual neuron can only signal the overall motion in its receptive field and not the behavior of individual features. Although the visual system could still recover information about individual features through a careful combination of cells with overlapping receptive field it is interesting to note that the overall activity in the population of neurons already represents a smoothed velocity field that is not keeping track of individual features per se in agreement with what our experiments 1 and 2 suggest.

In summary, we have presented experiments and perceptual demonstrations that strongly support the use of a surface interpolation scheme in the extraction of structure from motion in the human visual system. The companion paper (Hildreth et al 1991) will present a computational implementation of such a scheme based on Ullman's Incremental Rigidity Scheme (1984). This implementation will use velocity based information as suggested by our psychophysical findings.

References

Ando, H. (1992) Dynamic reconstruction and integration of 3-D structure information. Ph.D. thesis in preparation

Craton, L. & Yonas, A. (in press) Kinetic occlusion: Further studies of the boundary flow cue. Perception & Psychophysics.

Dosher, B. A., Landy, M. S. & Sperling, G. (1989) Kinetic depth effect and optic flow I. 3D shape from Fourier motion. Vision Research. 29, 1789-1813.

Husain, M., Treue, S. & Andersen, R. A. (1989) Surface interpolation in 3-D structure-from-motion perception. Neural Computation. 1, 324-333.

Johansson, G. (1975) Visual motion perception. Scientific American. 232, 76-88.

Miles, W. R. (1931) Movement interpretations of the sihouette of a revolving fan. American Journal of Psychology. 43, 392-405.

Mitchison, G. J. & McKee, S. P. (1985) Interpolation in stereoscopic matching. Nature. 315, 402-404.

Mitchison, G. J. & McKee, S. P. (1987) Interpolation and the detection of fine structure in stereoscopic matching. Vision Research. 27, 295-302.

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.

Nakayam, K., Shimojo, S. (1990) Towards a neural understanding of visual surface representation. in The Brain, Cold Spring Harbor Symposium on Quantitative Biology. Sejnowski, E. R., Kandel, C. F., Stevens, C. F., Watson, J. D. (eds.) Cold Spring Harbor Laboratory: New York. 55: 911-924.

Nakayama, K. & Tyler, C. W. (1981) Psychophysical isolation of movement sensitivity by removal of familiar position cues. Vision Research. 21, 427-433.

Norman, J. F. & Todd, J. T. (1992) The visual perception of affine stretching transformations. Perception & Psychophysics. in press,

Ramachandran, V. S., Cobb, S. & Rogers-Ramachandran, D. (1988) Perception of 3-D structure from motion: The role of velocity gradients and segmentation boundaries. Perception & Psychophysics. 44, 390-393.

Royden, C., Baker, J. & Allman, J. (1988) Perception of depth elicited by occluded and shearing motions of random dots. Perception. 17, 289-296.

Saidpour, A., Braunstein, M. L. & Hoffman, D. D. (1992) Interpolation in structure from motion. Perception & Psychophysics. 51, 105-117.

Shimojo, S., Silverman, G. H. & Nakayama, K. (1989) Occlusion and the solution to the aperture problem. Vision Research. 29, 619-626.

Snowden, R. J., Treue, S., Erickson, R. E. & Andersen, R. A. (1991) The response of area MT and V1 neurons to transparent motion. Journal of Neuroscience. 11, 2768-2785.

Thompson, W. B., Kersten, D. & Knecht, W. R. (1992) Structure-from-motion based on information at surface boundaries. Biological Cybernetics. 66, 327-333.

Todd, J. T. & Bressan, P. (1990) The perception of 3-dimensional affine stucture from minimal apparent motion sequences. Perception & Psychophysics. 48, 419-430.

Treue, S., Husain, M. & Andersen, R. (1991) Human perception of structure from motion. Vision Research. 31, 59-75.

Ullman, S. (1984) Maximizing rigidity: The incremental recovery of 3-D structure from rigid and nonrigid motion. Perception. 13, 255-274.

Würger, S. M. & Landy, M. S. (1989) Depth interpolation with sparse disparity cues. Perception. 18, 39-54.

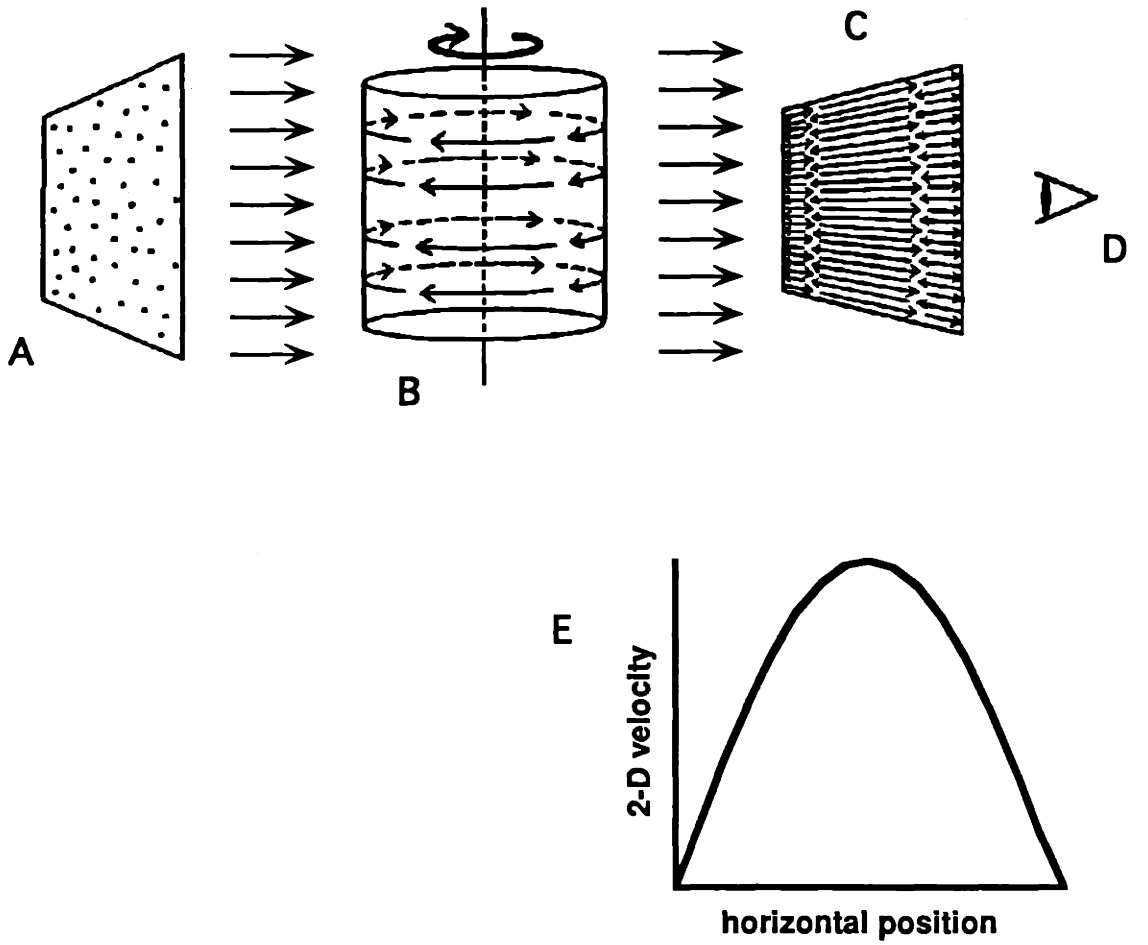


Figure 3.1

Cartoon of our stimulus generation algorithm
Points are plotted randomly on a 2-D square (A). They are then projected onto a cylinder (B) which is rotated. The projection of the rotating cylinder forms the visual display (C) that the subject observes (D). The velocity varies as a half-sinusoid across the display with the highest speeds int the center.

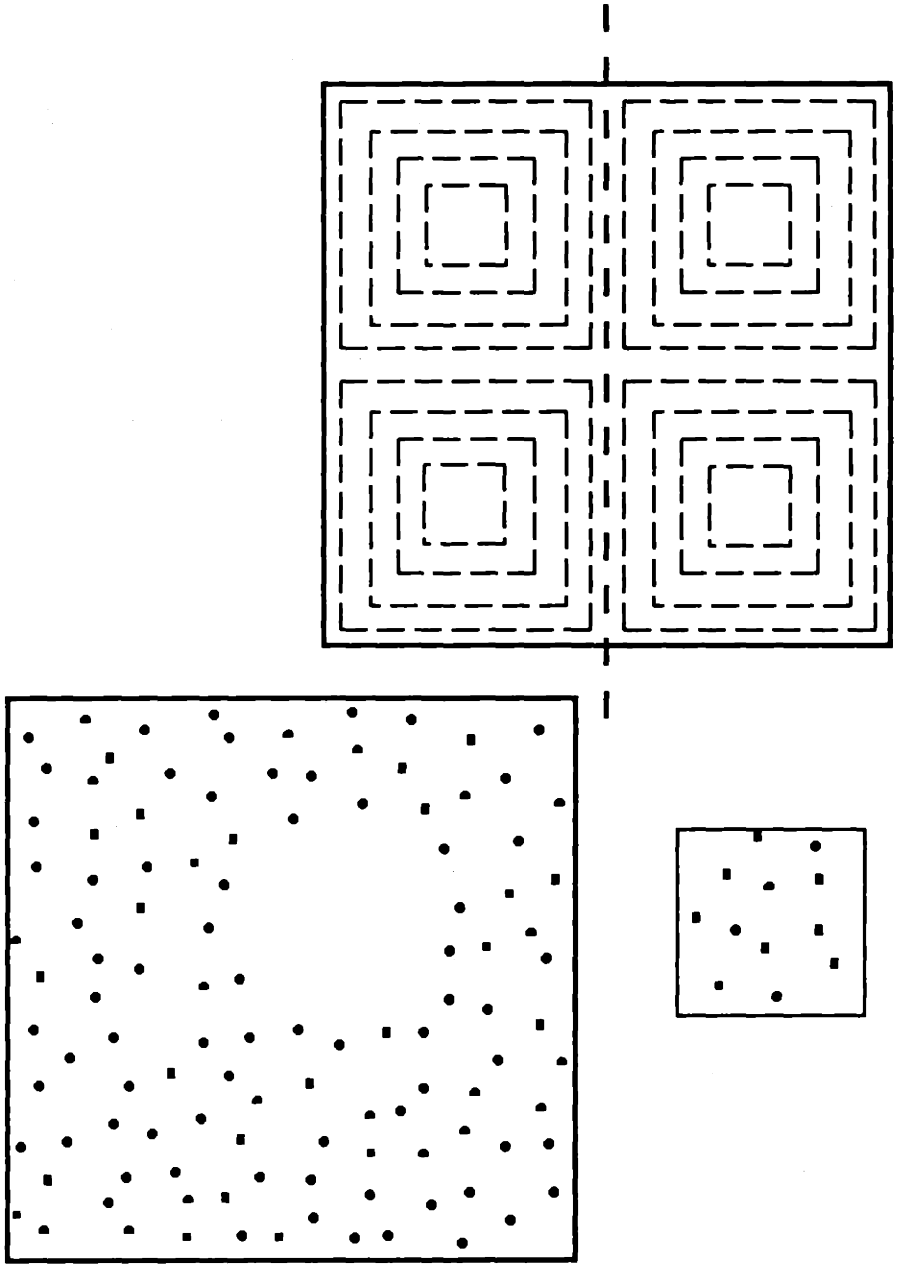


Figure 3.2

Mask positions for experiment 1
(A): To-scale drawing of the vsarious mask sizes and positions used in experiment 1.
(B): Exmple of the effect of using the mask drawn with a dotted line in A on on surface of our stimulus. the cut out stimulus part is represented in C. This mask size represented the threshold of performance in experiment 1.

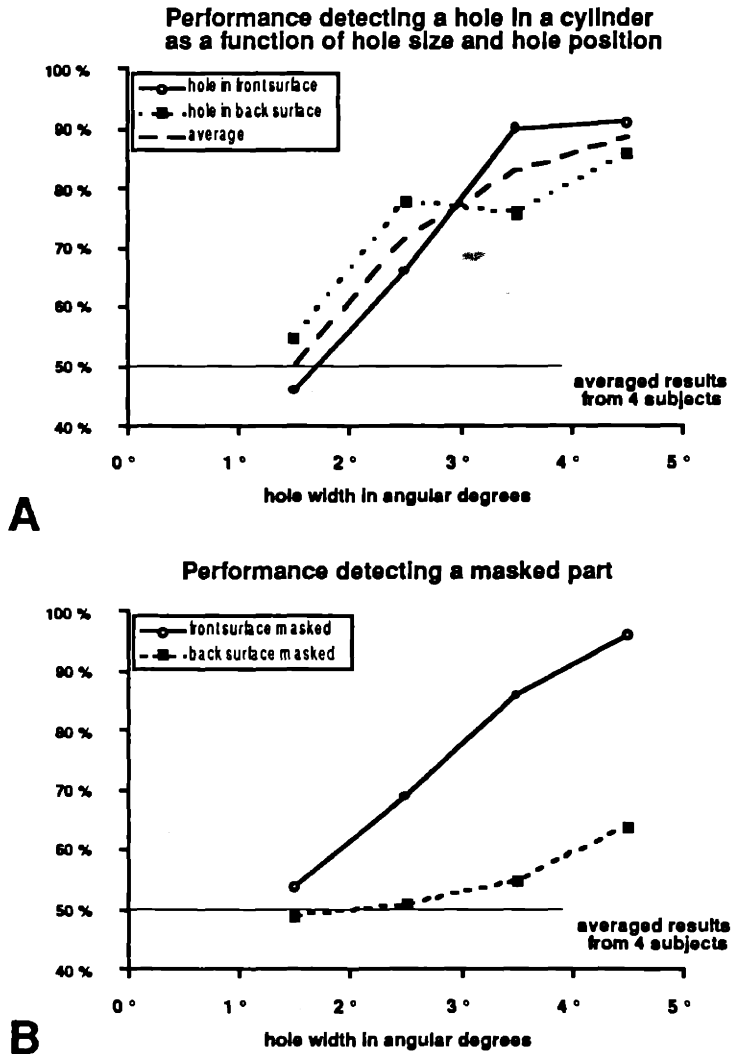


Figure 3.3

Experiment 1:
(A): Performance detecting a hole in a cylinder as a function of hole size and hole position.
(B): Performance detecting a masked part.

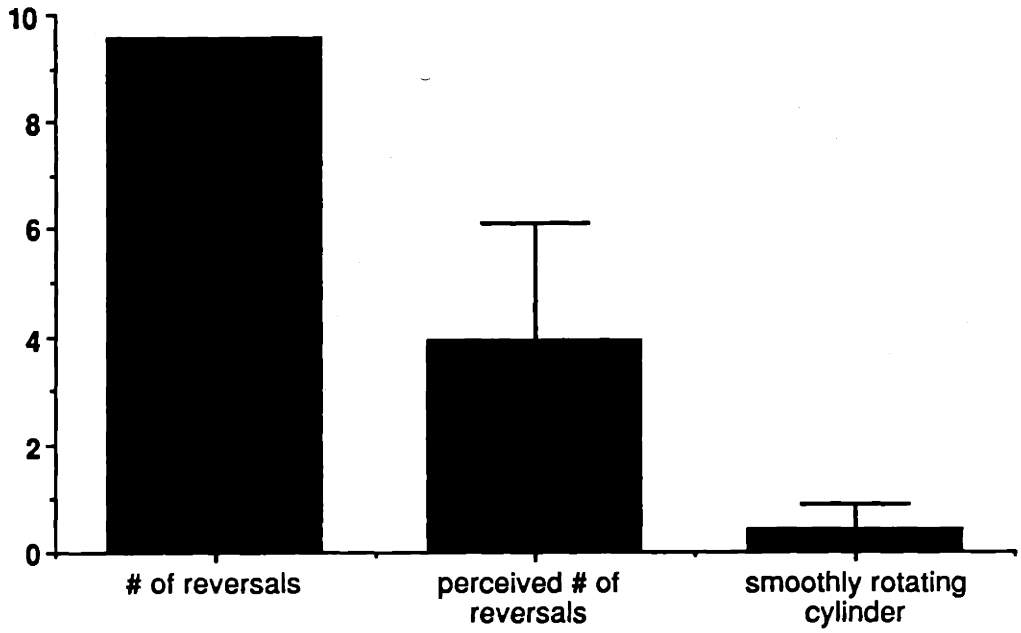


Figure 3.4

Experiment 2
left column: Number of changes of direction in the stimulus
center column: Number of perceived changes of direction of rotation of the cylinder
right column: Number of perceibed changes in direction of roattion using a smoothly rotating cylinder.

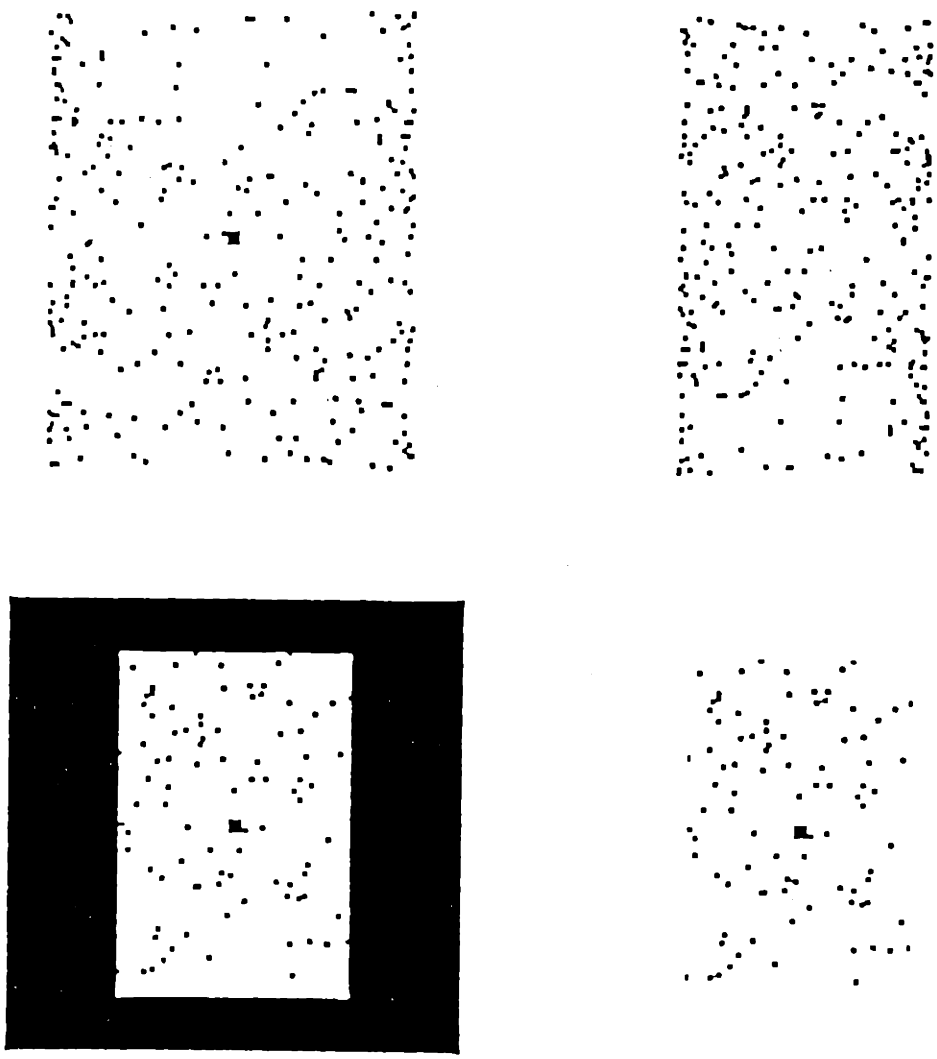


Figure 3.5

Individual frames from the displays investigating the role of surface boundaries.

A: Complete cylinder.
B: Complete cylinder of smaller diameter.
C: Cylinder masked by visible mask (extrinsic border).
D: Cylinder masked by invisible mask (intrinsic border).

Chapter 4

Looking ahead: The Analysis of Velocity Gradients in Area MT

Introduction

Velocity gradients are ubiquitous in our environment. They play two important roles in visual perception. In large optical flow fields they contain important information about observer heading while for smaller stimuli they determine the perceived 3-D structure of moving objects in the perception of structure from motion.

The analysis of optical flow fields has received a much attention from psychophysics (Gibson, 1950; Regan et al., 1982, 1985, 1986; Regan, 1986; Warren et al., 1988a, 1988b) and computational vision (Koenderink, 1986; Longuet-Higgins and Prazdny, 1980; Koenderink and Van Doorn, 1981; Rieger and Lawton, 1985) as well as electrophysiology (Duffy and Wurtz, 1991a, 1991b; Saito et al., 1986; Tanaka et al., 1986, 1989a, 1989b; Andersen, Snowden, Treue, & Graziano, 1990). The current understanding holds that area MST in primate visual cortex, with its large receptive fields and sensitivities to rotating, expanding and contracting patterns, is critically involved in the analysis of optical flow. Several studies have demonstrated position invariance[1] in the response properties of these neurons (Saito et al., 1986; Andersen, Graziano, & Snowden, 1990; Duffy and Wurtz, 1991a; Lagae et al., 1991). This shows that the receptive fields cannot simply be understood as a mosaic of simple direction tuned subunits. It is not clear though how the sensitivities of the MST neurons originate. Two alternatives are possible: the tuning of MST cells could result from the careful combination of inputs with relatively small receptive fields that are responsive to patterns moving in a particular direction and with a specific speed.

1 Position invariance means that a neuron that is tuned for example to counterclockwise rotation will retain this tuning even at stimulus locations far enough apart to locally stimulate the receptive field with opposite direction.

Alternatively the inputs themselves could show a preference for velocity gradients, albeit simpler than the ones preferred by the MST cells.

Such inputs could come from area MT where most of the input received by area MST originates. Besides constituting the building blocks for MST receptive fields, MT cells could analyze the small velocity gradients that exist in structure from motion stimuli. The role of velocity gradients in the perception of structure from motion is well understood from psychophysical and computational points of view as I demonstrated in the two preceding chapters. Electrophysiologists on the other hand have so far not investigated this issue. Thus we are trying to establish whether area MT neurons are sensitive to velocity gradients and could thus serve an important role in structure from motion perception and the early stages of optical flow field analysis. Several lines of argument point to area MT as a prime candidate for such a role. Earlier work already established that lesions to area MT will selectively affect structure from motion perception (Siegel and Andersen, 1986). The size of receptive fields in area MT seems well matched to the rather small stimuli commonly used in studies of structure from motion while the large receptive fields common in MST seem to lack the fine resolution that the perception of structure from motion requires. In fact computational studies suggest that gradient-selective cells must have relatively small receptive fields to detect surface curvature (Longuet-Higgins & Prazdny, 1980; Koenderink, 1986). Furthermore psychophysical studies by Nakayama and colleagues have suggested the presence of cells sensitive to gradients of motion which have receptive field sizes similar to MT neurons (Nakayama, Silverman, MacLeod, & Mulligan, 1985). Finally, MT might be the area where receptive field properties provide a link between the simple direction and

velocity tuned cells of area V1 and the complex structure of MST receptive fields.

The visual environment contains an infinite variety of different velocity gradients. Since this is the first physiological study using such stimuli, we limit ourselves to stimuli whose gradients represent simple linear increases or decreases of local velocity along any line that crosses the stimulus[2]. These velocity gradients all belong to the same class and thus can be represented easily in what we call a "deformation space" (see below). The deformation space has great resemblance to the idea of a "spiral space" developed in our laboratory as a way to characterize the response properties of MST cells to rotating, expanding and contracting stimuli.

Methods

A detailed description of our recording methods has appeared elsewhere (Snowden et al., 1991) and this section will therefore be limited to a brief overview and a detailed description of the stimuli used.

Two male rhesus monkeys were trained to fixate a small fixation point, while ignoring any other stimuli, and to signal the dimming of the fixation point by releasing a key. Using a scleral search coil technique (Robinson 1963) the animals' eye movements and point of fixation were closely monitored. Visual stimulation was provided to the receptive field of individual neurons during this 4 — 6 sec period of fixation. Electrode penetrations were made through a chamber implanted over the

2 Obvious other possible stimuli are for example the half sinusoidal velocity profiles used in our structure from motion stimuli investigated in detail in the two preceding chapters.

parietal cortex in one monkey and over area V1 in the other monkey. The electrode's position within the chamber, the depth of recording as well as the properties of the cells encountered during the penetrations were used as an aide in determining whether encountered cells were indeed in area MT. No histology is yet available since the animals are still used in experiments.

Experimental protocol

Stimuli were presented on a large HP CRT screen at a viewing distance of 57 cm. They consisted of random patterns of bright dots (30 ft. lambert) upon a dark background. Each dot was approximately 1 mm in diameter, and thus subtended about 6 min arc. We choose the CRT monitor rather than the video monitor used in some of the experiments described in chapters 5 and 6 since it offers a higher resolution. In our set-up we have 200 addressable locations within a centimeter on the screen rather than the 35 that we can achieve with a regular video monitor.

The rate of screen refresh was 50 Hz. Each trial commenced with the onset of the fixation point. After 1 sec a stimulus appeared if the animal had pulled the key and was successfully fixating. This stimulus was extinguished after 1 sec (i.e. 50 frames), and another stimulus appeared for 1 sec after a 1 sec delay. The fixation point dimmed 0.2 — 2.0 secs after the end of the last stimulus; thus a complete trial lasted 4.2 — 6 secs. In this manner we were able to present two stimuli per trial.

Stimuli

All our stimuli (except for the light bars used for initial mapping of the receptive fields) were random dot patterns. Introduced into vision research over 30 years ago (MacKay, 1957, 1961; Held & White, 1959; Julesz, 1960) they allow for prolonged stimulation without resorting to the repetitive sweeps often needed when using bars as stimuli. Random dot patterns also allow the generation of complex visual stimuli while maintaining a high level of control over stimulus variables like contrast, stimulus size, stimulus shape, etc..

To determine the preferred direction of motion for a cell we presented random dot patterns moving in eight different directions spaced 45° apart. These patterns moved behind a square virtual aperture. For each direction of motion tested the aperture was oriented so that the sides of the aperture were either oriented parallel or orthogonal to the direction of motion of the pattern. The computer algorithm that generated the movies would wrap any points that would otherwise have crossed the sides around to the opposite side of the stimulus. The preferred direction of the neuron determined in this way was used throughout the rest of the test performed on this cell[3].

After determining the cells preferred direction we establish its best speed by presenting it with patterns moving in that direction at speeds between 1 and 128°/sec.

To determine if a cell is gradient tuned we present it with various velocity gradients all moving in the cell's preferred direction. The average speed of each pattern is equal to the preferred speed of the cell as determined in the previous test. The velocity of every point in the pattern varies as a function of

[3] It should be noted that since all our stimuli (except the ones used to determine the cell's preferred direction) are moving in the preferred direction of the cell we use the terms velocity and speed interchangeably throughout this chapter.

its position within the gradient. We used two different functions.

(1) Shear stimuli

In these stimuli the velocity gradient is oriented perpendicular to the direction of motion. Thus the velocity of a particular point will not change while it is moving across the stimulus but neighboring points will have different velocities. All gradients have positive or negative linear slopes. We arbitrarily define patterns in which the velocity *increases* from the right of the stimulus to the left (when facing in the direction of motion) as having negative slopes and call them counter-clockwise (CCW) shearing stimuli[4] (Fig. 4.1A). Correspondingly we define those in which the velocity increases from left to right as having positive slopes and call them clockwise shearing stimuli (Fig. 4.1B). For these experiments we used two different slopes. The steeper velocity gradient would start at zero °/s on the slower end of the display and would reach twice the preferred speed of the cell under study at the opposite end. The shallower gradient would start at half the preferred speed and reach 1.5 times the preferred speed. Note that for both stimuli the average speed as well as the speed in the center of the stimulus is equal to the preferred speed of the cell.

(2) Compressive/Stretching stimuli

In these stimuli velocity varies *along* the direction of motion. Thus a point in a compressive gradient will *decrease* its velocity while crossing the display while always having the same velocity as its neighbors (Fig. 4.1D). Correspondingly points in stretching gradients will *accelerate* while crossing the

[4] We chose this nomenclature since in a counterclockwise *rotating* dot field velocity also increases from right to left when facing in the direction of motion. Note though that all the dots in the shearing stimuli move along a straight path.

display (Fig. 4.1C). As for the shear gradients we used two different slopes, one ranging from zero °/sec to twice the preferred velocity and one ranging from half the preferred velocity to 1.5 times the preferred velocity.

As mentioned above points that moved across the sides of the flat velocity profiles used for determining the preferred direction and speed of a cell were simply wrapped around to the opposite side of the stimulus. This simple technique is sufficient to insure equal dot density across the stimulus for these simple patterns as well as our shear gradients.

The acceleration or deceleration of individual dots in our compression stimuli on the other hand would lead to changes in stimulus density across the display especially for those compressive gradients in which dot speeds decrease all the way to zero at one end of the display. We therefore employed two techniques to eliminate this density cue in our displays.

(a) Special dot wrap-around. Expansive dot gradients that start with speeds of zero at one end and with even density distributions across them, will remain evenly distributed, but the density will be continuously falling due to the fact that the distance of any two points will constantly increase even though they are moving in the same direction[5]. We prevent the density in our displays from falling by replotting any points that cross the stimulus boundaries back into the stimulus. Notice that this replotting has to be done randomly across the stimulus rather than using the wrapping-around method employed for the flat velocity profiles. This replotting requires a more elaborate approach when the lowest speed in the stimulus is larger than zero. We generate those stimuli by first generating a larger stimulus whose gradient starts at a speed of zero and then

5 This phenomenon is well known in astronomy where the finding that any two stars are moving away from each other lends support to the idea of an expanding universe.

masking this stimulus to show only a smaller extent of the velocity gradient on the stimulus display.

Compressive gradients are generated by first computing a stretching gradient moving in the opposite direction and then reversing the order of the individual frames making up the stimulus. Notice that in the resulting stimuli dots will disappear while approaching the zero speed stimulus edge. Thus there is no "piling-up" of dots at that edge.

(b) Limited dot lifetimes. Replotting dots within our stretching stimuli and removing dots in the compressive stimuli generates transient events that could possibly influence the response of cells to these patterns. To insure that this transiency does not influence our findings we introduce it into all of our stimuli. This is achieved by using dots of limited lifetimes. The dots move along a continuous path for only a short period of time, their lifetime. After its lifetime a dot would be randomly replotted within the stimulus. We used a lifetime of 300 ms (15 frames) for all our random dot patterns which was long enough to not substantially affect the percept of motion while at the same time providing a significant amount of transiency, masking the transiency generated by the appearing and disappearing dots in the stretching and contractive gradient stimuli respectively.

As mentioned in the introduction our stretching, compressive and shearing dot patterns are members of a continuous family of stimuli which only vary along one circular dimension[6]. This dimension is the angle between the direction of the vector describing the velocity gradient (the "gradient vector") and the direction of dot motion in the pattern. If the gradient vector points in the direction of stimulus motion the stimulus is stretching. If the gradient

[6] This is true only if all the gradients have the same steepness of slope.

vector points in the opposite direction the stimulus is being compressed, while gradient vectors orthogonal to the direction of pattern motion occur in shearing stimuli. Figure 4.2 depicts how these stimuli can be represented in a continuous fashion. The four stimulus types described above (stretching, compression, clockwise shear, counter clockwise shear) form the cardinal axis of a coordinate system that we term "deformation space". But as also shown in Figure 4.2 there are stimuli that fall between these cardinal directions. These stimuli combine elements of stretching or compression with shear components.

Since this deformation space represents the one dimension along which the stimuli we use in these experiments vary, one might expect that cells that are selectively responsive to one or more of these stimuli are indeed tuned for the direction of the gradient vector. In the result section we will thus use the coordinate system described in Figure 4.2 to plot our results.

Results

Since these experiments are still ongoing this section will only deal with preliminary results gained from individual cells.

Figure 4.3a shows the speed tuning curve of a MT cell. The cell's preferred speed was 4°/s. That was therefore the base speed used for our gradient tuning test. The results are plotted in Figure 4.3b. The four open circles forming a square in the 'deformation' space represent the response of the neuron to the flat velocity profile moving at 4°/s while the four corners formed by the thick lines represent the response to the four different

velocity gradients. As can be seen the neuron responds the same to three of the gradients as to the flat velocity profile while the clockwise shear stimulus elicited a stronger response. The difference in responses is statistically significant (see Figure 4.3b legend for details).

Figure 4.4a shows the speed tuning curve of another MT cell. The cell was equally responsive to a pattern moving at 4 or 8°/sec in its preferred direction (represented by the arrow in the figure) while showing an inverted tuning curve to the anti-preferred direction. We therefore conducted our test for gradient tuning using an average speed of 4 as well as of 8 °/sec. The results are plotted in Figure 4.4b. The graph shows that both a compressive stimulus and a clockwise shearing stimulus are as effective as the flat velocity profile in driving this cell. Both the stretching and the counter-clockwise shear on the other hand drive the cell better. The response differences are highly statistically significant (see Figure 4.4b legend for details).

The similarity between the response evoked by the stretching and CCW shear suggest that the cell is actually most responsive to an intermediate direction for the gradient vector in 'deformation' space. Since recording the two cells presented in Figure 4.3 and 4.4 we have increased the numbers of stimuli types used in this study to the eight types depicted in Figure 4.2.

Discussion

Although the differences in response between the best flat velocity profile and the preferred stimulus in the 'deformation space' introduced here are small[7], they are statistically significant. Further our stimuli are so well controlled that no other cues than the velocity gradients themselves can account for the differences in neural response.

Thus the preliminary data presented above indicate that there are neurons in area MT of the primate that can encode more complex stimulus properties than direction and speed of simple moving patterns. Some MT cells are more strongly activated by velocity gradients than by flat velocity profiles moving in their preferred direction and speed. This implicates these cells in the perception of structure from motion and makes them well suited as input neurons to MST cells.

7 It is not surprising that the differences in neural response rate are small given that the flat velocity profile was optimized to generate the highest response achievable with a non-gradient stimulus.

References

Allman, J., Miezin, F. & McGuiness, E. (1985) Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). Perception. 14, 105-126.

Andersen, R. A., Graziano, M. & Snowden, R. (1990) Translational invariance and attentional modulation of MST cells. Society for Neuroscience Abstracts. 16, 7.

Andersen, R. A., Snowden, R. J., Treue, S. & Graziano, M. (1990) Hierarchical processing of motion in the visual cortex of monkey. Cold Spring Harbor Symposia on Quantitative Biology. LV, 1990.

Duffy, C. J. & Wurtz, R. H. (1991) Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. Journal of Neuroscience. 11, 1329-.

Duffy, C. J. & Wurtz, R. H. (1991) Sensitivity of MST neurons to optic flow stimuli. II. Mechanisms of response selectivity revealed by small-field stimuli. Journal of Neuroscience. 11, 1346.

Gibson, J. J. (1950) The perception of the visual world. Houghton Mifflin, Boston.

Golomb, B., Andersen, R. A., Nakayama, K., MacLeod, D. I. A. & Wong, A. (1985) Visual thresholds for shearing motion in monkey and man. Vision Research. 25, 813-820.

Held, R. & White, B. (1959) Sensory deprivation and visual speed: An analysis. Science. 130, 860-861.

Koenderink, J. J. (1986) Optic flow. Vision Research. 26, 161-180.

Lagae, L., Xiao, D., Raiguel, S., Maes, H. & Orban, G. (1991) Position invariance of optic flow component selectivity differentiates monkey MST and FST cells from MT cells. Investigative Ophthalmology and Visual Science (Supplement). 32, 823.

Longuet-Higgins, H. C. & Prazdny, K. (1980) The interpretation of a moving retinal image. Proceedings of the Royal Society, London. 208 B, 385-397.

MacKay, D. M. (1961) Visual effects of non-redundant stimulation. Nature. 192, 739-740.

Nakayama, K., Silverman, G., MacLeod, D. I. A. & Mulligan, J. (1985) Sensitivity to shearing and compressive motion in random dots. Perception. 14, 97-241.

Regan, D. (1985) Visual flow and direction of locomotion. Science. 227, 1064-1065.

Regan, D. & Beverley, K. I. (1981) How do we avoid confounding the direction we are looking and the direction we are moving? Science. 215, 194-196.

Regan, D. & Beverly, K. I. (1985) Visual responses to vorticity and the neural analysis of optic flow. Journal of the Optical Society of America. 2, 280-283.

Regan, D., Erkelens, C. J. & Collewijn, H. (1986) Necessary conditions for the perception of motion in depth. Investigative Ophthalmology and Visual Science. 27, 584-596.

Rieger, J. H. & Lawton, D. T. (1985) Processing differential image motion. Journal of the Optical Society of America. A 2, 354-360.

Rogers, B. J. & Graham, M. (1979) Motion parallax as an independent cue for depth perception. Perception. 8, 125-134.

Rogers, B. & Graham, M. (1982) Similarities between motion parallax and stereopsis in human depth perception. Vision Research. 22, 261-270.

Roy, J. P., Komatsu, H. & Wurtz, R. H. (1992) Disparity sensitivity of neurons in monkey extrastriate area MST. Journal of Neuroscience. 12, 2478-2492.

Saito, H., Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y. & Iwai, E. (1986) Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. Journal of Neuroscience. 6, 145-157.

Siegel, R. M. & Andersen, R. A. (1986) Motion perceptual deficits following ibotenic acid lesions of the middle temporal area in the behaving rhesus monkey. Society for Neuroscience Abstracts. 12, 1183.

Tanaka, K., Fukada, Y. & Saito, H. (1989) Underlying mechanisms of the response specificity of expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. Journal of Neurophysiology. 62, 642-656.

Tanaka, K., Hikosaka, K., Saito, H., Yukie, M., Fukada, Y. & Iwai, E. (1986) Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. Journal of Neuroscience. 6, 134-144.

Tanaka, K. & Saito, H. (1989) Analysis of motion of the visual field by direction, expansion/contraction, and rotation cells clustered in the dorsal part of the medial superior temporal area of the macaque monkey. Journal of Neurophysiology. 62, 626-641.

Warren, W. H. & Hannon, D. J. (1988) Direction of self-motion is perceived from optical flow. Nature. 336, 162-163.

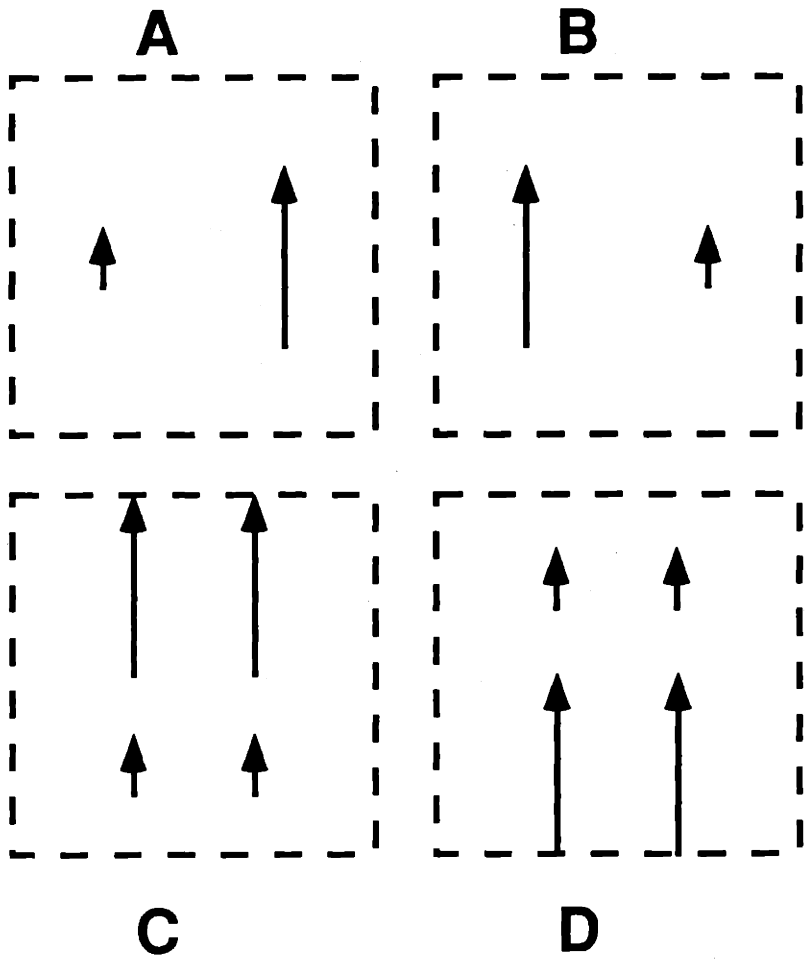


Figure 4.1

(A) Counter-clockwise shear
(B) Clock-wise shear
(C) Stretching
(D) Compression

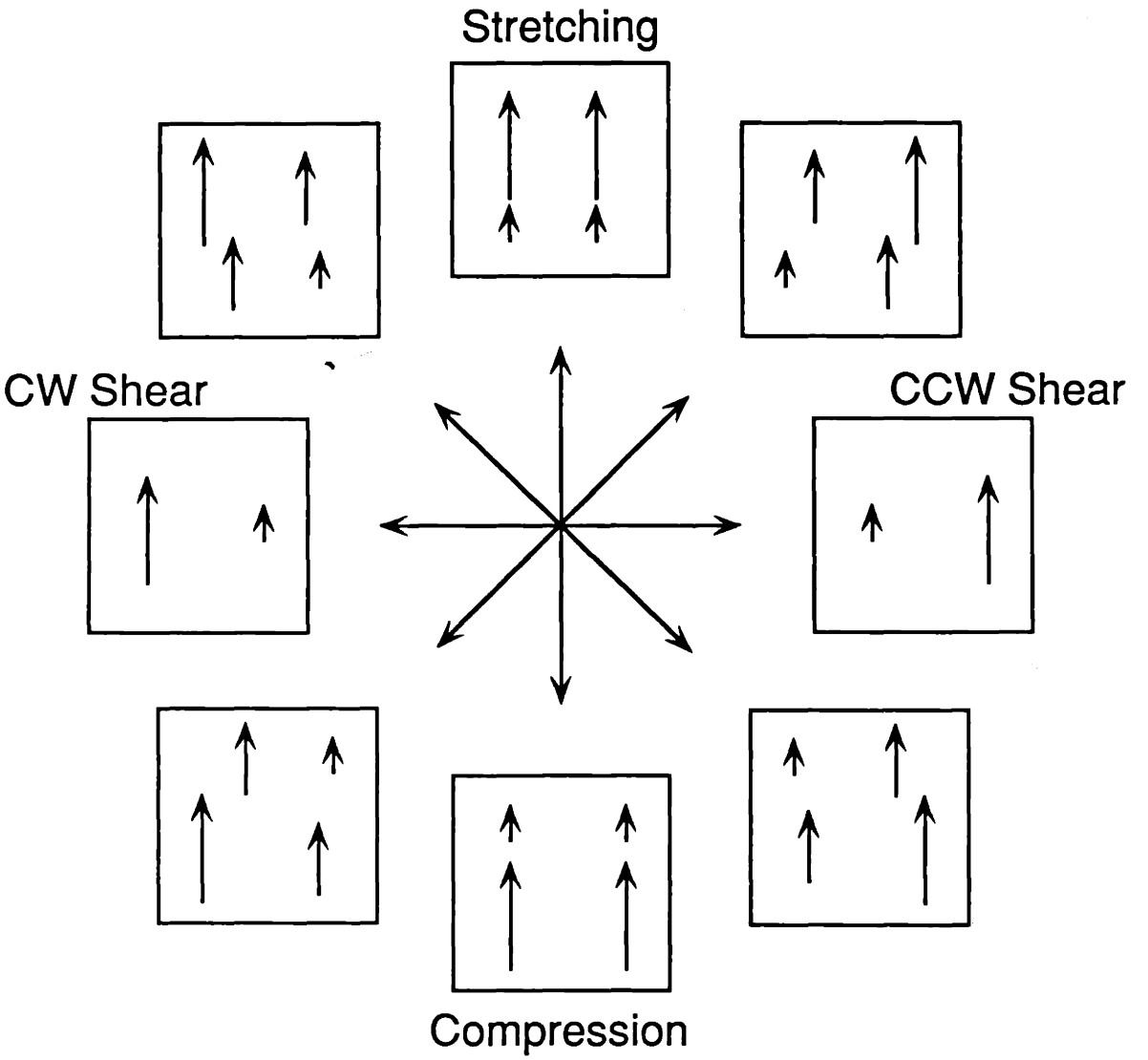


Figure 4.2

All gradient stimulus types used in this experiment at their appropriate positions in polar coordinates in the 'deformation space' introduced in the text.

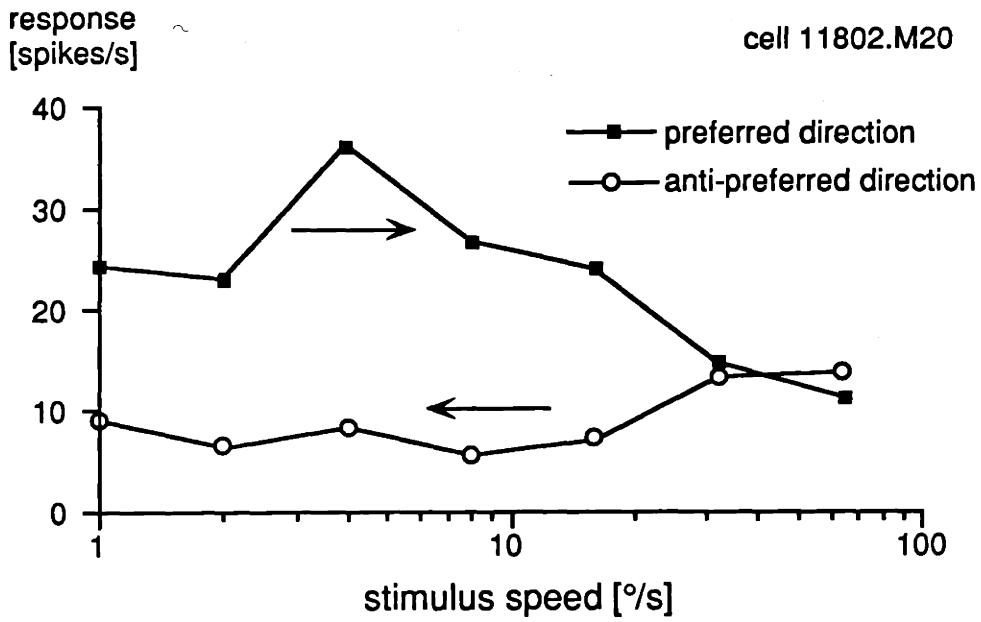


Figure 4.3a

Speed tuning curve for a MT neuron. The stimulus was moving either in the preferred direction (rightwards) or anti-preferred direction (leftwards). Speeds ranged from 1 to 64 °/sec and are plotted here on a logarithmic axis. The preferred speed of this cell is 4 °/s.

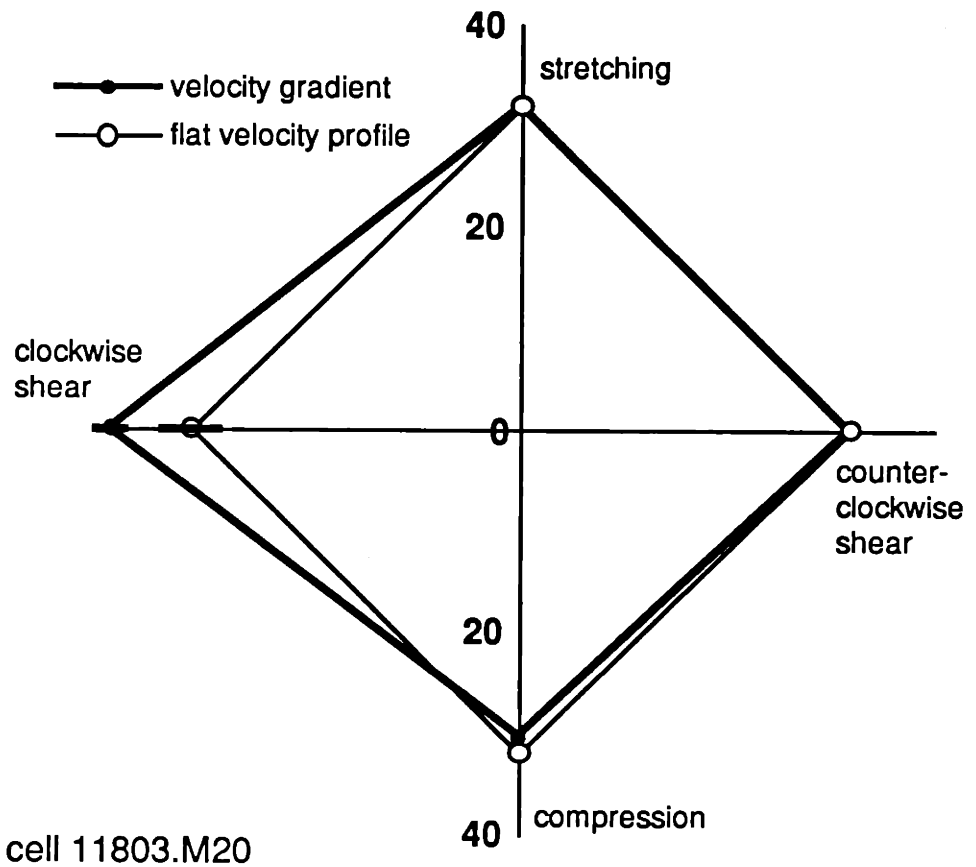


Figure 4.3b

Response of a MT neuron to a random dot pattern moving in the cell's preferred direction and at the cell's preferred speed (open circles) and the same cell's response to 4 different velocity gradients (see Figure 4.1). Notice that this cell shows an increased response to the clockwise shearing stimulus. The lines through the clockwise shear response and the response to the flat profile represent standard errors of the mean. The clockwise shear response is larger than the flat response at .018 probability (as determined by paired 2-tailed t-test). The numbers alongside the vertical axis refer to the cell's response in spikes per second.

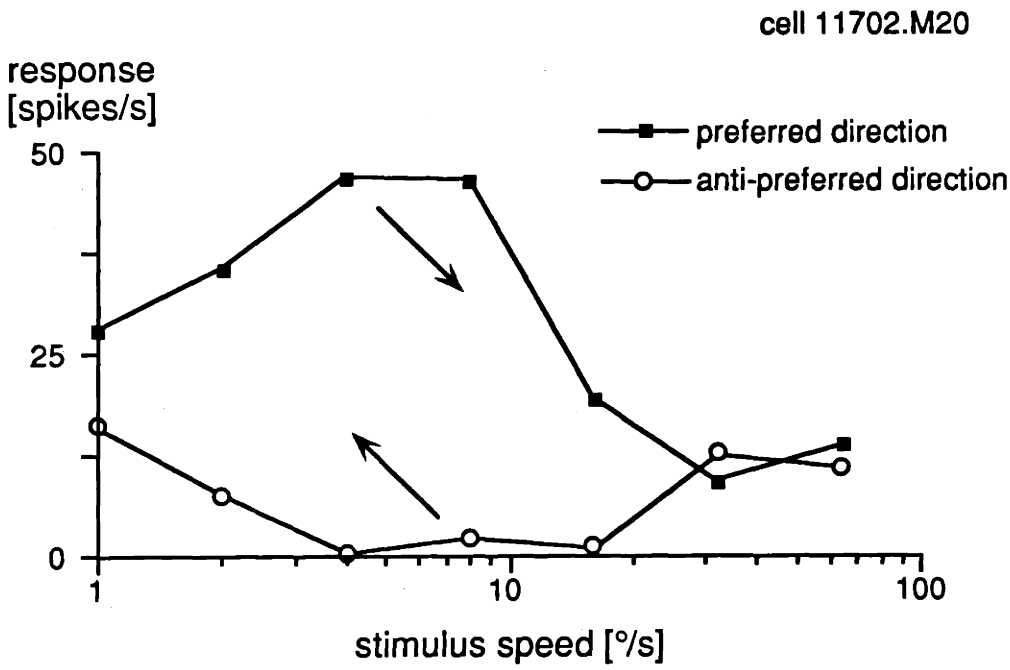


Figure 4.4a

Speed tuning curve for a MT neuron. The stimulus was moving either in the preferred direction (down to the right) or anti-preferred direction (left up). Speeds ranged from 1 to 64 °/sec and are plotted here on a logarithmic axis. The preferred speed of this cell was between 4 and 8 °/s.

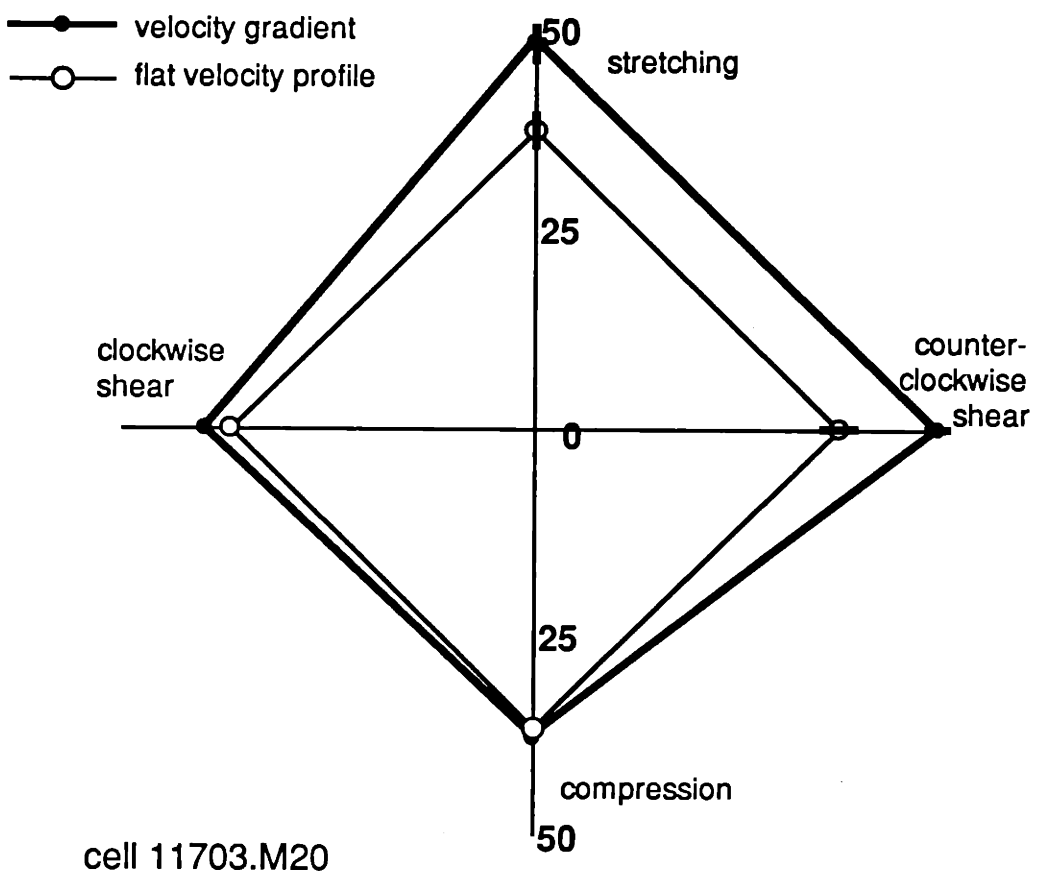


Figure 4.4b

Response of a MT neuron to a random dot pattern moving in the cell's preferred direction and at the cell's preferred speed (open circles) and the same cell's response to 4 different velocity gradients (see Figure 4.1). Notice that this cell shows an increased response to both stretching and counter clockwise shearing stimuli suggesting that its preferred stimulus in this 'deformation' space is a stimulus combining stretching as well as counter-clockwise shearing components. Lines through the response circles represent standard errors of the mean. The counter clockwise shear response and the expansion response are larger than the flat response at .0007 and .0045 probability respectively (as determined by paired 2-tailed t-tests).

Direction Selectivity and Receptive Field Structure in Macaque Cortical Areas V1 and MT

Chapter 5

The Response of Area MT and V1 Neurons to Transparent Motion

Introduction

Motion transparency exists whenever two different motions occur at the same local region in an image. It is quite common in natural images, being found under a variety of movement conditions. An obvious example of motion transparency is the view one receives while looking through the window of a moving vehicle in the rain with water streaking down the glass. A less obvious, but perhaps more common instance of transparency occurs when a shadow moves across a textured background. If only a single motion vector is allowed at each local region in the image then one would either perceive the shadow border dragging the texture along with it, or the shadow border being rendered stationary by the texture. Specular reflections represent another potential transparent condition during movement. For instance, when a person wearing glasses rotates his head the specular reflections remain stationary, yet we do not perceive the head moving and the spectacles remaining still. Even motion discontinuities or "borders", which are generated by object motion or observer motion, are a type of motion transparency. In the local region of the motion border, there is the differential motion of the object and the surround.

Computer algorithms developed to analyse moving images have considerable difficulty with motion transparency (Fennema and Thompson, 1979; Horn and Schunck, 1981; Heeger, 1987; Yuille and Grzywacz, 1988; Bülthoff et al., 1989; Wang et al., 1989). Each involves calculating the local motion components and integrating them, to smooth or average the local velocity field. This computation is typically performed to improve signal to noise ratios, interpolate motion across areas of the image where there is sparse data, and to solve the aperture problem.

An unfortunate result of this computation is that every point in the image can only have a single motion vector assigned to it, and as a result these algorithms are blind to motion transparency. For instance, under transparent conditions where there is motion of the same speed but in opposite directions, these models would report no motion at all. The fact that artificial systems have such difficulty analysing transparency, and that its occurrence is so frequent in natural scenes, suggests that the human visual system has developed specialized methods for perceiving it.

Motion transparency also presents difficulties for current models of motion analysis by the primate visual system. The prevalent model for motion-direction selectivity employs inhibitory interactions between groups of neurons (Barlow and Levick, 1965). Such inhibitory interactions would suppress motion selective cells in transparent situations, and render the visual system blind to it, much like the computer systems mentioned above. On the other hand, facilitatory mechanisms for direction selectivity, originally discussed by Barlow and Levick (1965) and observed in several studies (e.g. Ganz and Felder, 1984), would be unaffected by transparent stimuli. At a higher stage than initial measurement of motion direction, the two influential models of motion perception of Reichardt (1961) and Adelson and Bergen (1985) both involve subtracting the outputs of oppositely tuned direction detectors, which would give a zero output to opposite, transparent motions.

The ability to sort out the different motion components in transparent moving stimuli extends to the occulomotor system. Kowler and her colleagues (1984) have demonstrated that subjects are able to perfectly pursuit a moving textured field in the presence of a superimposed stationary field and vice versa.

Also the ability to see transparent motion presumably extends to sub-human primates. The rhesus monkey has been shown to have similar motion processing capacities to humans, in particular both species can perform similarly in reacting to a 3 dimensional rotating hollow cylinder merely through the 2 dimensional projection of the cylinder (Siegel and Andersen, 1988). As this involves 'transparent motion' it is therefore most likely that these monkeys can also see transparent motion and thus provide a suitable animal in which to explore the mechanisms which lead to this percept.

In the primate brain motion information appears to be processed in a hierarchical manner. Neurons early in the visual system respond well to moving objects but do not show a differential response to movements in different directions. The first location to do so is the striate cortex, area V1, (Hubel and Wiesel, 1968). In this area around one third of the cells have a directional response, and these cells seem to be concentrated in the upper sublayers of layer 4 (4B, 4Ca, 4A) and layer 6 (Hawken et al., 1988). A strong projection leads from both layer 4B (Maunsell and van Essen, 1983; Shipp and Zeki, 1985) and layer 6 (Fries et al., 1985) to area MT (or V5) whose neurons are almost exclusively directional (Zeki, 1974; Albright, 1984; Mikami et al., 1986a). Damage to this extrastriate area compromises performance on visual motion tasks, but spares other visual functions such as contrast sensitivity to stationary gratings (Siegel and Andersen, 1986; Newsome & Paré, 1988). Such a hierarchy of projections suggests that each area may play a different role in motion perception (e.g. Maunsell and Van Essen 1983; Movshon et al., 1985; Tanaka et al., 1986; Saito et al., 1986; Andersen and Siegel 1990; Andersen, et al., 1990a), with area MT elaborating the information provided by V1.

To gain insights into how a biological system processes complex visual environments which contain transparent motions we have recorded the response of neurons from both area V1 and MT in the alert monkey to plain motions, transparent motions and motion boundaries. We found the cells in area V1 to respond well to their preferred direction of motion even in the transparent condition, whereas area MT neurons were substantially inhibited under the same transparent conditions. These results suggest that the primate visual system solves the transparency problem by allowing, in area V1, more than one motion vector to be represented at each local region in the image. As a result subpopulations of V1 neurons are tuned to different directions of motion at the same retinal location, perhaps representing the early stage for segmenting different, transparent surfaces. In area MT, the inhibitory interaction of opposed motions may contribute to the smoothing or averaging of the velocity field that is a feature of both the artificial motion analysis systems and models of the primate visual system mentioned above, and used to reduce noise and interpolate surfaces from moving features. The differential processing of direction information by areas V1 and MT provide further evidence for a hierarchical system of motion analysis. (Preliminary versions of the results presented here have appeared elsewhere (Erickson et al., 1989; Snowden et al., 1990)).

Materials and methods

Preparation of animals

Two male rhesus monkeys (Macaca Mulatta) were used. The animals were trained to fixate in a dimming-detection reaction-time task with the head immobilized. Each trial was initiated with the illumination of a light emitting diode (LED) which the animal was required to fixate after pulling back upon a lever within 800 ms. After a randomized period of between 3 and 4 seconds the LED dimmed, and if the animal released the lever within a period 150 - 600 msec after the dimming he received a drop of apple juice reward and the next trial took place after an interval of approximately 5 sec. The animals eye positions were monitored during the fixation period by the magnetic search coil technique (Robinson, 1963; Judge et al., 1980) and if the animals eye moved at a speed above 15 degrees/sec the trial was terminated without reward. Eye position was monitored every 35 ms and the standard deviation of eye position on the successful trials was less than 9 min arc for each of the animals (mean = 6.1 mins arc). Details of the training and surgical procedures have been published previously (Andersen and Mountcastle 1983; Golomb et al., 1985; Andersen et al., 1990b; Andersen et al., 1990c).

Stimuli

The animal sat in a primate chair placed 57 cm from a monitor. The experimental room was dimly lit (0.01 lux) and screen luminance was also 0.01 lux. The screen was surrounded with black cardboard which could be used for projecting handheld stimuli. The stimuli used for quantitative analysis were all random dot patterns undergoing apparent motion. For informal testing and receptive field mapping computer

generated and hand held bars were also used. Random dot patterns were produced by plotting points at random locations within a square area. To produce movement the coordinates of each point in the next frame were suitably changed. Any point that now fell outside the square window was wrapped to the opposite side. Thus the effect was of a sheet of random dots moving behind a stationary window. The size of the area, number of dots, speed and direction of movement were all under experimental control. In addition each dot was displayed only for a limited lifetime (500 ms) and was randomly repositioned within the pattern upon their death. The dots were given a random starting lifetime so that points died in an asynchronous manner (Morgan and Ward, 1980). In addition it was possible to make patterns in which the dots were stationary but had limited lifetimes. In the experiments using a 'stationary noise' pattern the stimulus consisted of dots which did not die over the course of the stimulus presentation, while the 'dynamic noise' was produced by stationary dots of lifetime 48 msec.

These 'movies' were produced in advance of the experiments and stored in the memory of the computer. The stimuli were generated and displayed using two different systems at different times. During our early recordings (which constitute around 44% of the V1 cells and 65% of the MT cells) the stimuli were generated by software running on a PDP 11/73 and displayed upon a large Hewlett Packard oscilloscope (P31 phosphor) with a nominal frame rate of 35 Hz (i.e. each frame was displayed for 28 ms before the next frame was shown). Each stimulus was displayed from stimulus onset to the time of the lever release. The dots were 1700 lux with a diameter of approximately 1 mm. During later experiments stimuli were generated via a Number Nine graphics board housed in a AST 386 computer and displayed on a raster display (NEC multisync XL) running at 60

Hz (each frame therefore lasted 16.6 sec). Here the stimulus was divided into three periods - 1 sec stimulus, 1 sec blank, 1 sec stimulus, hence two stimuli could be presented on each trial which greatly enhanced the speed of data collection. The dots were of similar luminance and size as in the previous setup. No obvious differences have been observed between these different setups.

The stimuli were tailored to some extent to match the properties of the cells being recorded. During the early recordings we used a smaller field size for area V1 (radius 1.5 deg, 64 dots per surface = 9.2 dots/deg^2) than MT (radius 3.0 deg, 64 dots per surface = 2.3 dots/deg^2) in an attempt to compensate for the smaller receptive fields in V1. However, as we kept the number of dots in the stimulus constant this in itself means that the dot density is different for these two patterns. In control experiments (see results section) we concluded that dot density (or number of dots in receptive field) has little effect per se, and so in the later parts of the experiments we used identical stimuli for both V1 and MT (radius 1.5 deg, 64 dots per surface = 9.2 dots/deg^2).

Recording Procedure

Recordings were made with glass coated platinum-iridium electrodes advanced through the intact dura. The chamber was placed over area V1 such that area MT was accessible with deep penetrations through V1. After each neuron was isolated an attempt was made to drive the cell by presenting bars and/or dots over various parts of the screen and monitoring activity through an audio feedback, while the animal performed the fixation task. Once the receptive field had been determined the stimulus was placed in its center. Then series of blocks of trials

were presented until all experiments had been completed or until the cell was lost. A block of trials typically consisted of a set of stimuli whose ability to drive the cell we wished to compare. For example, the first test usually run consisted of a series of 5 movies which contained 8 directions of motion, a stationary pattern and a blank interval for assessing the spontaneous rate. These 5 movies were presented in a pseudorandom interleaved order until the monkey had completed 6 - 10 trials for each movie successfully. The results were then quickly inspected to determine the preferred direction of motion, and the next block of trials chosen in accord with these results. The tests performed varied depending upon the nature of the cell and the current experimental needs. Typically up to 10 blocks of trials could by run on a single cell (approximately 600 trials) and this could take around 2 hours to complete. Many cells were lost before all tests could be complete therefore the tests were run in a hierarchical fashion depending upon our needs.

Identification of cells and areas

Cells were initially assigned to a visual area based upon their functional properties, receptive field position and size, their position relative to other cells and the depth along a penetration. Over a series of penetrations the receptive field positions and sizes of the first cells encountered were recorded and the position of V1 was calculated from the maps published by Dow et al. (1985). Our population of V1 cells was recorded at eccentricities of 0.5 - 3 deg near the vertical meridian. Area MT was identified using physiological criteria including direction selectivity of nearly all cells, receptive field size, and topographic organization (i.e. Maunsell and van Essen, 1983; Newsome and Paré, 1988; Gattass and Gross, 1981). Our

population of MT cells was recorded at eccentricities of 1 to 10 deg.

During the last few penetrations in one monkey we laid down marking lesions at some recording sites, by passing DC current through the tip of the electrode. The animal was later sacrificed by an overdose of sodium pentobarbital and perfused transcardially with heparinized saline followed by formalin. Four guide wires were also inserted using the coordinate system of the microdrive shortly after the animal was sacrificed. These were then used as markers for reconstructing where the older electrode penetrations had occurred. The animal's brain was sectioned in the horizontal plane every 40 mm and every sixth sections stained with thionin for cytoarchitecture and its neighboring section stained for myelin by the method of Gallyas (1979). MT and V1 were identified on anatomical grounds and the sites of lesions identified and late electrode tracks reconstructed (see figure 5.2). This enabled us to verify the position of MT with respect to our recording sites. The second animal is alive and is currently engaged in other experiments.

Data recording and analysis

Cell discharges were digitised and their times recorded by the computer for off-line analysis. The times of all other trial events (e.g. stimulus onset, key down, reaction time etc.) were also stored by the computer. Post-stimulus time histograms were then constructed by collating the data from the same trial types into 50 ms bins and averaging over the number of trials. Examples of such histograms along with the raw spike trains from each trial are shown in figure 5.3. For quantitative analysis a 0.8 sec interval was chosen (commencing shortly after the response latency) and the mean firing rate and standard deviation within this time period was calculated.

Our total sample of V1 (N = 167) and MT (N = 87) cells was screened for reliability of response. To be included in our analysis we had to be able to drive the cell (cells which we could only inhibit were excluded) and we had to be able to get reliable differences between two conditions (e.g. direction of motion). To assess this reliability we performed a t-test between the two conditions A and B via the formula:

$$A_{mean} - B_{mean} \; / \; ((A_{SD}^2 / N_A) + (B_{SD}^2 / N_B))^{1/2} \qquad (1)$$

where SD stands for the standard deviation and N for the number of trials respectively. Usually we did not perform this test between every two conditions; rather we chose the response in preferred, antipreferred and spontaneous rate conditions and did the comparisons between these. Any cell that had a score > 5.0 on any test was judged to be significantly effected by our stimuli and was included in the rest of the analysis (V1, N = 130; MT, N = 72).

Results

Response to transparent motion

1) Individual cells

Initially the direction and speed tuning of each cell was tested by handheld stimuli and by quantitative tests involving interleaved trials where motions in four or eight different directions were presented. The preferred direction of motion was the direction that elicited the greatest mean firing rate and the antipreferred direction was 180 degrees from the preferred direction. The response to single surfaces moving in the preferred direction, antipreferred direction and a combination of the former two stimuli superimposed was then assessed in another block of interleaved trials. This latter stimulus appeared as two transparent sheets of dots drifting through one another and we therefore termed it the two-surface stimulus. The response of two representative neurons to these stimuli are shown in figure 5.3. The upper part of each section of the figure shows the spike trains elicited on each trial, while the lower part is the post-stimulus time histogram averaged over these trials. The left-hand column of this figure depicts the response to the preferred direction alone, the middle section the response to the antipreferred direction alone, and the right-hand column the response to the two-surface stimulus. The time of the stimulus presentation is indicated by the bar below each histogram. Inspection of these individual examples shows that these cells have quite different responses to different directions of motions. In order to quantify this difference we calculated an index of directionality (I_d) between opposite directions of motion:

$$I_d = 1 - A / P \qquad (2)$$

where P is the response to the preferred stimulus and A the

response to the antipreferred stimulus, both after the spontaneous rate (measured by interleaved trials where the animal simply fixated the LED and no stimulus was presented) was subtracted. An I_d of 0.0 indicates no difference between the two directions, while one of 1.0 indicates no response in the antipreferred direction. Values greater than 1.0 indicate that antipreferred stimulus reduced the level of activity in the neuron to below its spontaneous level. The I_d values of the cells in figure 5.3 are given in the figure legend.

The response to the two-surface stimulus is presented in the far right of figure 5.3. It is notable that in the case of the V1 neuron, illustrated in the top half of the figure, the response to the two-surface stimulus was similar to the response to the preferred surface alone. On the other hand, the response of the MT neuron, illustrated in the bottom half of the figure, was considerably less than for the preferred surface. To quantitatively compare the two surface and preferred direction responses, we calculated a 'suppression index' (I_s) in a manner similar to the one used for the direction index:

$$Is = 1 - \alpha / P \qquad (3)$$

where α is the response to the two-surface stimulus (again after the spontaneous rate has been removed). Values less than 0.0 indicate the response was greater to the two-surface rather than the preferred stimulus, 0.0 means they were equal, and values greater than 0.0 indicate the preferred stimulus gave a greater response than the two-surface stimulus. The I_s values of the cells show in figure 5.3 are given in the legend.

2) Population data

For each neuron on which we performed this experiment a I_s was calculated, and these are presented in terms of their frequency of occurrence for each of our two populations in figure

5.4. The cells from V1 show a strong tendency to have a lower I_s than those from MT. The median I_s for V1 is 0.04 (indicating the response to the two-surface stimulus was similar to the response to the preferred stimulus) with cells falling on either side of the 0.0 mark (indicating suppression in some cells and facilitation in others). The median Is for MT was 0.54 with some cells being inhibited below their spontaneous rate ($I_s > 1.0$) by the two-surface stimulus. No cells from MT had a $I_s < 0.0$ indicating that all cells were suppressed below the rate for the preferred stimulus alone. The difference between the two populations' data was confirmed by an independent samples t-test ($t=5.7$, $p<0.0001$).

The proportion of directionally selective responses ($I_d > 0.7$) was 30% in area V1 and 92% in area MT. This is in agreement with previous estimates (Zeki, 1974; De Valois et al., 1982; Albright, 1984; Mikami et al., 1986a; Hawken et al., 1988). As MT shows both higher I_ds and I_ss than V1, it is possible that these indexes may stem from a common cause - that of increased inhibition in the antipreferred direction. Thus we plotted the I_d against the I_s in figure 5.5 for each neuron where this information was obtained. The upper section plots the data from area V1, and the lower section data from area MT. It can be seen that in the area of overlap of direction indices in V1 and MT, which was greatest between about .6 and 1.2, V1 I_ss still showed considerably lower suppression indices than area MT. Thus, the difference in suppression indices between areas V1 and MT cannot be explained solely on the basis of greater directional indices in MT. A comparison of indices was also made within areas. For the V1 cells we found no significant correlation between these indexes (Spearman rank correlation = 0.14, $p > 0.2$). For the MT cells we did find a significant

correlation (Spearman rank correlation = 0.45, $p < 0.001$). This indicates that, within area MT, the I_s was related to the I_d in a manner that the greater the I_d the greater the I_s. However, it is also noticeable that many cells which were not inhibited below their spontaneous rate by the antipreferred direction alone still had I_ss > 0.0 indicating that this stimulus has an excitatory response presented alone, but an inhibitory effect when presented in conjunction with the preferred direction. This presumably reflects an inhibition by the anti-preferred motion which is masked by an excitatory but unspecific response to high-contrast random dot patterns (also see our discussion of division-like inhibition below).

Possible role of dot density

We were concerned that our results might be influenced by the fact that our two-surface stimulus normally contained twice as many dots as the single surface-stimulus (as a result of superimposing two single-surface stimuli to form the two surface stimulus). If we assume that increasing the dot density would increase the response then we might underestimate the suppression effect. We thus performed a control experiment on cells in both V1 and MT in which we determined the suppression index for a two-surface stimulus made up from two surfaces of either 32 or 64 points. In both cases the response of the cell to the preferred direction alone was determined with a single surface stimulus containing 64 points. The data from V1 and MT is portrayed in separate plots (figure 5.6). As can be seen lowering the density of the two-surface stimulus to equal the density of the single-surface stimulus had no systematic effect on the I_s. The most likely reason for that result is the rapid saturation of the cells' response with increasing dot density (see figure 5.12). For single surface stimuli the saturation usually occurs with just a few dots.

In comparing data from MT and V1 it should be noted that the receptive field sizes at any given eccentricity are very different with area MT having larger receptive fields (Gattass and Gross, 1981). Since we used stimuli of the same dot density in area V1 and MT the number of dots in the receptive fields of the two areas was quite different. To make sure that this would not systematically effect our I_s measurements we obtained responses from 8 cells in MT and 12 cells in V1 to dot densities ranging from 0.45 to 28.8 dots/deg^2 in each surface - a 64 fold change in density. Again no systematic change was found in either area, though it became noticeable that the data obtained from V1 cells at the low densities became erratic and the correlation with the values obtained at a higher dot density was poorer. This result is perhaps not surprising since at the lowest dot density for the V1 cells, there could exist intervals where no dots were actually in the receptive field. These controls show that as long as at least modest dot densities are used this parameter does not have a substantial effect on the data.

Thus the addition of the antipreferred direction of movement to the preferred direction of movement causes a suppression of response in many cells. This suppression is apparent in all MT cells and is far stronger than that found in V1.

Effects of stationary and dynamic noise

The above results show that movement in the antipreferred direction can suppress activity driven by the preferred direction. However, it is possible that the suppression might have nothing to do with movement per se, but that it is a more general effect .

We tested this prediction on 13 MT cells by replacing the antipreferred direction with either a stationary pattern or a

dynamic noise pattern (a series of uncorrelated frames: see methods) in the combination stimulus. The response to each stimulus was scaled relative to the response to the preferred stimulus alone, and the mean and standard error are shown in figure 5.7. Once again the antipreferred direction caused a large suppression. Both the stationary pattern and dynamic noise also caused some suppression (of similar magnitude), but this was mild compared to the antipreferred movement. These differences were confirmed by non-parametric statistics (Mann-Whitney U: 1) antipreferred and stationary pattern, $p < 0.005$; 2) antipreferred and dynamic noise, $p < 0.01$; 3) stationary and dynamic noise $p > 0.5$). These results show that stationary pattern or dynamic noise do reduce the response compared to the preferred direction alone, but this suppression is considerably smaller than that produced by the coherent, antipreferred movement.

Motion Borders

The two surface stimulus contains dots moving in opposite directions. These motion vectors are randomly positioned so to produce the impression of two overlapping transparent surfaces. It is of interest to know whether splitting the different motion vectors into discrete groups, so that the motion is no longer 'transparent' but now appears as a set of motion borders would also produce the same suppressive effect described above.

An example of this test is shown in figure 5.8 for an MT cell. Above each response histogram is a cartoon of the stimulus presented, and the lower parts show the spike rasters and the post-stimulus time histograms. Parts a & b of the figure show the response to the anti- and preferred direction respectively. Note that the dots were confined to three stripes within the circle, and that the stripes without dots for the leftward motion

are the ones with stripes for rightward motion (the cell was also tested with the alternate arrangement of motion direction and gave very similar results). This cell was direction selective and showed no response to its antipreferred stimulus. Part d of the figure shows the response to the super-imposition of the two stimuli (now alternate stripes were filled with dots moving in opposite directions). The cell gave little or no response to this stimulus, which was similar to its response to the transparent stimulus, shown in part c. Clearly then separating the dots into discrete stripes does not allow the cell to respond as it does to the preferred direction alone. We performed a comparison between overlapping (transparent) and segregated (striped) displays by calculating an Is (as above) for each of these displays on 17 MT neurons. Figure 5.9a plots these indexes against one another. As can be seen many of the points fall close to a line with a slope of 1.0 indicating little difference between the results obtained for the separated and transparent stimuli. However, there is a trend for the response to be greater to the separated stimulus (more points fall below than above the dashed line at 45°) indicating that there was less suppression than under the transparent conditions. We attempted to push this test further in 12 cells by presenting only two stripes (i.e. the upper half moving in one direction and the lower half in the opposite direction and vice versa). We often found that these two stimuli gave very different results (i.e. when the preferred was, for example, in the upper half the cell responded strongly, while when the preferred direction was in the lower half the cell gave a very poor response). We therefore averaged the response from the two conditions to compare with the transparent motion stimuli. The data are presented in figure 5.9b. The result is similar to that obtained with 6 stripes - the cell is still inhibited below the response to the preferred

direction alone but the suppression is not as strong as under the transparent conditions. Britten and Newsome (1990) have recently reported results very similar to these.

Direction tuning of suppressive effects

We were interested in relative degree of suppression generated by directions of motion other than opposite to the preferred direction. Sixty four dots were always drifted in the cells preferred direction, and another 64 dots drifted in one of eight directions each separated by 45 degrees. The results of a representative MT cell are shown in figure 5.10a. As can be noted directions other than those 180° from preferred also produced inhibitory effects in that the response to the two surface stimulus is less than that to the preferred stimulus alone. Figure 5.10b demonstrates the direction tuning curve of this neuron for single directions of movement. It is noticeable that the cell is strongly excited by movements 45° from the preferred direction. Yet this same movement (45° from preferred) causes the response to the preferred direction to be reduced! It is therefore clear that the same stimulus can cause an increase or decrease in the cells firing rate depending upon the context in which it is presented.

All of the 11 neurons on which we performed this test gave roughly similar results, though the width of the suppression tuning varied from cell to cell. In order to show this variation we normalized the response of each cell with respect to the firing rate in the preferred direction and calculated the mean and standard error of the tuning, these results are shown in figure 5.11. The suppressive effect develops quickly as direction of the suppressor is moved away from the preferred up to around 90°, and then increases very slowly as direction of the suppressor is increased to 180°. As can be seen from the

standard errors there was considerable variation from cell to cell in the shape of the suppression tuning curve.

The nature of the inhibition

A) Preferred/antipreferred titration

Inhibitory influences can modify the response of a cell in two ways (Blomfield, 1974). The first can be thought of as a linear mechanism where IPSPs and EPSPs simply add to produce the final total. This process produces a subtraction. The second process is a non-linear mechanism which produces a ratio of the excitatory and inhibitory input. This process produces a division. However, these can be regarded as the limiting cases and in reality strong hyperpolarizations or thresholds can serve to introduce non-linearities which can mimic a pure division-like process (for discussion see Amthor and Grzywacz, 1991)

In order to investigate whether the inhibitory effects described above are more like a subtraction or a division, we 'titrated' the effects of the preferred and antipreferred directions of movement by systematically varying the dot density in each of these directions. Two representative neurons from area MT are portrayed in figure 5.12a & b. The open symbols represent the neurons response to changes in the dot density of the preferred direction presented on its own. As dot density is increased the response of the neuron rises sharply then begins to saturate and asymptotes at fairly low dot densities. When in addition a number of dots were drifted in the antipreferred direction the shape of this function is altered. In each case the rise in response with dot density is less steep, and the greater the density of dots in the antipreferred direction, the more gradual the rise in the function. As the function saturates rapidly for the preferred direction alone there is a range of dot

densities which cause no increase in response to the preferred direction alone, but cause an increase in response when there are also some dots moving in the antipreferred direction.

An ideal subtractive inhibition would cause the function to be shifted by a constant amount down the vertical axis, and the size of this shift would be a monotonic function of the density in the antipreferred direction (figure 5.13a). Hence, the slope of this function should not be affected, which is clearly not the case (however, a subtractive inhibition followed by a non-linearity could produce a change in slope; Amthor and Grzywacz, 1991). A division-like inhibition does indeed predict that the slope of the function should become more gradual as dot density in the antipreferred direction increases, and therefore the results point to a division-like operation being at the root of the inhibitory effect. A simple division-like operation would predict that all functions would saturate at the same dot density, and because of the more gradual rise, at a different response rate (see figure 5.13b). We found, however, that the functions continued to rise after the dot density at which the preferred function alone saturated. These functions carried on rising with increasing dot density and saturated only when they reached the same response rate at which the function for the preferred direction alone saturated (which is at a much higher dot density). This implies that the mechanism of saturation occurs after the division-like operation to produce functions similar to those portrayed in figure 5.12. Such an operation can be conceptualized as a gain control upon the incoming excitatory information, which we reproduce by scaling of the effects of dot density in figure 5.13c. These functions were produced by dividing the horizontal axis by a constant (it can be conceptualized as stretching the function along the horizontal axis). In all MT neurons where we were able to obtain sufficient

results to distinguish between the various functions described above, all corresponded to those portrayed in figure 5.13c.

B) Direction tuning

A prediction of the division-like inhibition hypothesis is that the inhibition (when plotted on linear coordinates) increases with increasing response strength of the neuron. For an inhibition factor of 2, if a neuron would respond without inhibition at 10 spikes/sec then with inhibition it would respond at 5 spikes/sec; whereas a rate of 100 spikes/sec would be cut to 50 spikes/sec. One simple way to vary the response of MT neurons is to shift the direction of motion of the stimulus. If the inhibition is division-like we would expect to see the greatest net loss of spikes/sec when the movement is in the preferred direction and a decreasing loss as the direction is moved away from the preferred direction.

To test this prediction we measured directional tuning curves with and without 64 dots moving in the anti-preferred direction in an interleaved block of trials. Representative results are shown in figure 5.14. As can be seen the function with dots in the antipreferred direction are predicted by a simple division factor on the function obtained for the single surface. We also obtained results from the relatively rare V1 cells which showed suppression by using a higher dot density in the antipreferred direction; these too showed similar effects.

Discussion

Relationship to previous work

Most physiological investigations of visual function have concentrated upon the response of neurons to discrete small stimuli within their classical receptive fields. However, more recently these studies have been extended so that the test stimuli is placed in the context of a more complicated environment, which may extend well beyond the classical receptive field (review: Allman et al., 1985b). Allman et al. (1985a) have shown that for many MT cells of the owl monkey the response to a bar or dot pattern moving within the cells receptive field can be suppressed by the addition of a surround moving in the same direction, and some cells can be facilitated by the surround moving in a different direction. The current experiment differs from that of Allman et al. in that the two patterns are coextensive and the effects can be produced entirely within the classical receptive field. Hence, our result that the response to the preferred movement is suppressed by movements in a different direction (the opposite result to Allman et al.) does not contradict, but complements the results of Allman et al. (see also Tanaka et al. (1986) for similar results in area MT of macaques, and Hammond and MacKay (1981) and Hammond and Smith (1983) for similar results in cat striate cortex).

Hoffman and Distler (1990) recently reported presenting neurons in the nucleus of the optic tract in the pretectum and the dorsal terminal nucleus of the accessory optic tract with stimuli made by superimposing two oppositely moving random dot patterns. They report that the two neurons they presented with that pattern were firing less to the two-pattern stimulus than to either of the individual patterns alone, a result

comparable to what we find in area MT.

Transparency

The percept to the casual observer of our combination stimulus is that there are two transparent surfaces drifting through one another (Clarke, 1977; van Doorn et al., 1985). The formation of the percept of two surfaces must be achieved purely on the basis of the motion information available in the display, as all other cues such as colour and form have been deliberately removed.

a) Computational models

Much recent effort has been applied in attempting to produce algorithms which calculate the 'optic flow field' (Fennema and Thompson, 1979; Horn and Schunck, 1981; Hildreth, 1984; Heeger, 1987; Uras et al., 1988; Yuille and Grzywacz, 1988; Bülthoff et al., 1989; Wang et al., 1989). These algorithms follow the general principle of calculating a local velocity component at each point in the image (as in the primary motion detection stage) and then smoothing or averaging the resultant field (as in the motion integration stage) on the premise that the velocity in most parts of the image will be varying in a gradual manner (some algorithms perform these operations in discrete stages, while others do this simultaneously). These algorithms account well for some known properties of human motion detection e.g. motion capture (MacKay, 1961) and cooperativity (Chang and Julesz, 1985; Williams et al., 1986) but have difficulty where velocity changes sharply, such as at the juncture between objects (though some models have suggested solutions to this problem; Hutchinson et al., 1988; Grzywacz and Yuille, 1990). These models cannot handle transparent motion stimuli as a smoothing or averaging operation demands a single answer at each point in the image.

Indeed, with our opposed motion stimuli the smoothing operation would result in no computed motion whatsoever.

These models are not the only ones which fail under conditions of motion transparency. For instance, the models of Reichardt (1961) and Adelson and Bergen (1985) both involve a stage where motions of opposite direction are subtracted from one another (see Adelson and Bergen, 1985 for details) thus eliminating any response to our transparent stimuli. Some models (e.g. Watson and Ahumada, 1985) employ an OR gate between opposite directions, again this excludes motion transparency.

These computational algorithms work in two stages 1) measuring the local velocity components, 2) spatially integrating these measurements. It is therefore tempting to consider the hypothesis that the two areas we recorded from (V1 and MT) are the physiological counterparts of these stages. The relative lack of suppression in V1 suggests that these cells appear to act like directional filters, extracting their preferred movement from the two surface stimulus. Thus within V1 there would be a group of neurons signalling one motion, and another signalling the other direction of motion. As the interactions between these groups are weak they could be active at the same time and represent the two directions of motion present in the stimulus. This segmentation of the two directions of motion into two populations allows more than one local motion vector to be represented at each point in space, something current computational models do not allow at their output stage. Secondly, this segmentation might allow smoothing or averaging operations to be applied separately to each of the populations, hence providing a mechanism for representing optic flow fields

which are not inherently unique at each point in the image. Such a technique has been recently shown to enhance an algorithm for computing the structure-from-motion of a transparent rotating cylinder (Ando et al., 1990).

The cells within area MT, on the other hand, produced a far greater response to their preferred direction alone than when it was presented as part of a transparent motion stimulus. Hence, these neurons do not act as simple directional filters. The strong interactions seen in this area may be indicative of a spatial integration of the local velocity signals arising in V1, and as such MT is a strong candidate area for the spatial integration operations of the computational theories (Wang et al., 1989). It should be noted that MT could represent the subtraction stage proposed by Adelson and Bergen (1985) (as long as the early direction selectivity is set up using a facilitatory mechanism) if the result of the subtraction is greater than 0.

b) Psychophysics

From the above results we can make some clear predictions concerning the perception of transparent motions. If motion thresholds are governed by the actions of cells in V1 we should expect that single motion surfaces would be almost as detectable as when this motion is embedded in a transparent motion stimulus. However, much evidence points to MT being an area important for the perception of motion. Lesioning this area results in thresholds for detecting motion in random dot patterns being considerably elevated (Siegel and Andersen, 1986; Newsome and Paré, 1988), the action of single neurons seems well correlated with the perceptual thresholds of the monkey (Newsome et al., 1989), and many psychophysical motion thresholds are in accord with the know properties of this area (Golomb et al., 1985; Baker and Braddick, 1985;

Mikami et al., 1986b; Newsome et al., 1986; Snowden and Braddick, 1990). This predicts that a motion surface will be more detectable when presented on its own than when this motion surface is embedded in a transparent motion stimulus due to the inhibition prevalent within MT. Recent psychophysical evidence supports this idea by showing that movements in orthogonal directions reduce the upper displacement limit for apparent motion (Snowden, 1989; 1990). We found that the addition of stationary and dynamic noise to the preferred direction of motion reduces the response of the cell, but this reduction is not as great as that caused by the opposite direction of motion. This is in accord with psychophysical findings that show stationary and dynamic noise reduce subjects' ability to report direction of movement, but do not compromise this ability as much as coherent movement (Snowden, 1989). The upper displacement thresholds may be processed in MT, which has larger receptive fields compared to V1 and is consistent with greater inhibition in MT. Further, motions of opposite directions (as in the stimuli we used) can increase the observers contrast threshold for detecting movement (Mather and Moulden, 1983) and reduce the ability to see correlated movements (Lappin and Kottas, 1981) when compared to a single surface (see also Sutherland, 1961; Moulden and Mather, 1978; Watson et al., 1980; Stroymeyer, et al., 1984). All these authors interpret their results in terms of inhibitory interactions between different directions of motion, which we now suggest arise within area MT. In this context we should point out that even though MT neurons respond less to transparent motion than to motion in their preferred direction alone each of the two moving surfaces still activates a separate subpopulation of neurons (although the neurons' firing rates are on average reduced by 40 %).

Further psychophysical work has involved changes in the perceived direction of motion of a stimulus when two directions are superimposed (Marshak and Sekuler, 1979; Mather and Moulden, 1980), or when a single direction is viewed after prolonged inspection of a pattern moving in a somewhat different direction (Levinson and Sekuler, 1976; Mather, 1980). These authors also interpret their results in terms of inhibitory interactions between different directions of motion (Mather and Moulden, 1980).

Several studies of shape-from-motion perception using random dot stimuli indicate that the visual system can interpolate 3-D surfaces from sparse motion data (Siegel and Andersen, 1988; Husain et al., 1990; Treue et al., 1991; Siegel and Andersen, 1990). In these experiments the stimuli were transparent hollow rotating cylinders. A recent computational model of 3-D structure-from-motion perception which uses surface interpolation and accounts for the results of these studies, requires that the two surfaces present in the stimulus be segregated first and then smoothed (interpolated) individually for the structure-from-motion computation (Ando et al., 1990). Our data indicate that V1 would be able to perform such a surface segregation based on the opposite direction of motion of the front and back surface. MT, while dampened in its response by the transparent stimulus, might still be sufficiently activated to perform the smoothing and interpolation needed for the extraction of surfaces for perceiving structure from motion.

Figure 5.7 demonstrates that the addition of stationary and dynamic noise to the preferred direction of motion reduces the response of the cell, but this reduction is not as great as that caused by the opposite direction of motion. This is in accord

with psychophysical findings that show stationary and dynamic noise reduce subjects' ability to report direction of movement, but do not compromise this ability as much as coherent movement (Snowden, 1989). Such stimuli also activate many cells in areas V1 and MT. If, as suggested above, these cells which are activated then send inhibitory signals to cells which have different direction selectivity, then the overall effect upon a cell which is responding to its preferred direction of movement will be to reduce its response (Snowden, 1989). These results are therefore easily accounted for under the current theoretical framework. Also of great interest are recent reports that patients who have received damage to a brain area thought to be a human analog of area MT. The addition of small amounts of stationary or dynamic noise to otherwise coherently moving pattern completely masked the motion for these patients (Baker et al., 1990; Vaina et al., 1990). Similar effects have been observed in strobe-reared cats (Pasternak, 1990). This suggests that the inhibitory interactions which we demonstrate in area MT may play an important role in reducing noise.

The aperture problem

If a straight moving contour is viewed through a small aperture its direction of movement can only be estimated to within 180° (Wohlgemuth, 1911) - the so-called aperture problem. A possible solution to this problem is to have each oriented contour provide a 'line of constraint' for the possible motion, and thus two or more such contours will provide a solution to the aperture problem through their 'intersection of constraints' (Adelson and Movshon, 1982). Individual neurons could, in theory, respond to either the motion of each orientated contour (component motion) or to the overall movement of the pattern (pattern motion). Movshon et al. (1985) have examined this issue using plaid patterns (the superimposition of two

sinewave gratings). They found that all V1 cells respond in a component fashion, whereas a substantial number of MT neurons respond in a pattern fashion (see also Rodman and Albright, 1989). Movshon et al. (1985) interpret this result as suggesting a two-stage motion process - extraction of the motion perpendicular to the oriented contours, followed by a non-linear computation of the intersection of constraints (psychophysical evidence is also provided to support this claim). How this non-linear computation might be achieved is not suggested. Bülthoff et al. (1989) have, however, demonstrated that a computational model, involving a two-stage process not dissimilar to that suggested by the present data, simulates human performance when presented with the aperture problem. Their model is similar to the one suggested here in that it involves a local motion measurement followed by summation and competition between different direction of motion. Indeed they tentatively identify the first step with area V1 neurons and the summation and competition stages with area MT. The only major difference between their model and the current model is that the competition stage in their model is a 'winner-takes-all' whereas we suggest a division-like inhibition which does not necessarily produce a 'winner-takes-all' answer (though see Yuille and Grzywacz, 1989).

All cells we recorded from area MT showed a suppressive effect (see figures 5.4 & 5.5) and thus do not fall readily into two classes which could correspond to the component and pattern type responses of Movshon et al. However, from the data provided by Movshon et al. (1985) there appears to be a continuum rather than two discrete classes. The issue of the relationship between 'pattern versus component' and 'degree of suppression' therefore requires a more exhaustive and detailed

study before a statement can be made. It should be noted though that the basis for the distinction between pattern and component cells is the ambiguity inherent in plaids made from gratings because the aperture problem prevents the determination of the direction of motion of the single gratings. This is not the case for random dot patterns and thus two random dot patterns moving in different directions never cohere.

Motion segmentation

Psychophysical studies have demonstrated that objects can be segmented on the basis of motion information alone (Braddick, 1974). The large receptive fields of area MT suggests that boundaries defined by motion alone might appear blurry and ill-defined. However, perceptually these borders are very crisp. In the same vein, area MT neurons were shown to be suppressed by motion borders as they were by the transparent stimulus. Thus it would appear that the detection of motion borders are signalled by cells in another visual area. It is possible however that a lack of activity would be interpreted as a border, since the population image in area MT would be of cells signalling two directions of movement away from the border and much reduced activity would be present on the border. A similar suggestion was made by Grzywacz and Yuille (1990) on theoretical grounds. It is also possible that the non-classical surround of MT cells described by Allman et al. (1985a), which often facilitate responses from the classical receptive field when movement is in the opposite direction in the surround, could make MT sensitive to motion borders.

Models of direction selectivity

There are many models of how a neuron might achieve a difference in its response when the same pattern is moved in

different directions. One of the most influential models has been that put forward by Barlow and Levick (1965) after experiments on retinal ganglion cells of the rabbit. They found that flashing a single test bar within the receptive field elicited a response from the neuron, but the response to this test bar disappeared if another conditioning bar was presented immediately before and to one side of the test bar (thus mimicking the antipreferred direction of motion). If the conditioning bar was presented on the other side of the test bar (thus mimicking the preferred direction of motion) the response to the test bar did not disappear. They concluded that the directional response of the bar was achieved due to the first bar causing an inhibitory signal to be passed in one direction (the antipreferred), thus vetoing any response in this direction, while leaving the response to the other (preferred) direction unaffected. Such a model predicts the addition of the antipreferred direction onto the preferred direction (our two-surface stimulus) would cause many inhibitory signals to be generated which would reduce the response to the preferred direction. We found very few V1 cells which were affected in this way, and this raises the possibility that the mechanism whereby primate V1 neurons achieve directionality is not the inhibitory veto mechanism of the rabbit retinal. One possibility consistent with our findings in V1 would be a facilitatory mechanism, a mechanism considered by Barlow and Levick (1965) but which was inconsistent with the rabbit retina cells. Here the conditioning bar would not cause inhibition in the antipreferred direction, but would send a facilitatory signal in the preferred direction. If the Barlow and Levick model can be considered as a AND-NOT operation, the facilitatory model can

1 Gryzwacz and Amthor (1989) have argued that even in the rabbit retina direction selectivity is formed by a non-directional inhibition and a directional facilitation.

be considered as an AND operation (though an analogue version seems more likely). Such a mechanism would respond equally well to the preferred direction stimulus in the single and two surface conditions, just as many of the V1 neurons do.

The question of the mechanism of directionality has been considered many times in many species, and it is worth a brief review of the findings to put our hypothesis in context. Many models can be considered as variants of the Reichardt detector (Reichardt, 1961) developed from consideration of the fly's visual system. They have three qualities 1) two inputs separated in space 2) an asymmetry between the inputs so that different directions of motion cause different responses 3) an interaction to compare these two inputs (see Borst and Egelhaaf (1989) for a review). The Reichardt detector has a multiplication as its interaction, and multiplications can be achieved by either inhibitory or excitatory mechanisms (Torre and Poggio, 1978; Koch et al., 1986; Grzywacz and Koch, 1987). In cat striate cortex there is ample evidence for an inhibitory mechanism from experiments which are conceptually similar to that of Barlow and Levick (Goodwin and Henry, 1975; Goodwin et al., 1975; Ganz and Felder, 1984; Emerson et al., 1987; Baker and Cynader, 1988) and from pharmacological manipulation of the inputs to a neuron (see Sillito, 1979). However, other authors appear to find somewhat contradictory results. Douglas et al. (1988) fail to find any large conductance changes in area 17 neurons in cat which are predicted by the veto model (the so-called 'shunting inhibition'; Poggio and Koch, 1987). Movshon et al. (1978) repeated the Barlow/Levick two-bar interaction experiments on complex neurons of cat striate cortex and found evidence for facilitatory interactions rather than inhibitory ones. Hence, it appears that both facilitatory and inhibitory mechanisms exists in these species. While we have suggested

that many primate V1 neurons might use a facilitatory mechanism, it is clear that there may be some inhibitory mechanisms too. More direct experiments are required to reveal the relative importance of the two mechanisms. A possible reason for the use of facilitatory mechanisms, is that an animal with only inhibitory mechanisms would in principle not see transparent motions.

Barlow/Levick type experiments have also been performed in area MT of the alert macaque (Mikami et al., 1986a). They found evidence for inhibitory interactions in every neuron, and also for facilitatory interactions in most neurons. Our results are in accord with these findings. Do then the inhibitory interactions we find correspond to the workings of a Barlow/Levick type veto operation? We have reason to suspect not. The inhibition which is sent in the antipreferred direction of Barlow and Levick's model is caused, not by movement in the antipreferred direction, but just by the presence of a pattern at all. Hence, the addition of a stationary or dynamic noise pattern should reduce the response of MT neurons as effectively as movement in the antipreferred direction. We found this not to be the case. The addition of stationary or dynamic noise did reduce the response to the preferred direction somewhat but the reduction was considerably smaller than that elicited by the antipreferred direction. Secondly, if the inhibition of the Barlow/Levick model is confined to specific subunits within the receptive field (as suggested by Barlow and Levick (1965) and by Ganz and Felder (1984)) then splitting the motions into discrete sections should eliminate the suppressive effect. The suppressive effect was found to weaken under these conditions for the shear stimulus (see figures 5.8 & 5.9) but was certainly not eliminated.

If we reject the Barlow/Levick type inhibition as a candidate

for the suppressive effects shown in area MT what is the mechanism? One candidate mechanism is that the inhibition stems from competitive interactions between neurons with different preferred directions of motion, a notion with strong psychophysical (Levinson and Sekuler, 1976; Marshak and Sekuler, 1979; Mather and Moulden, 1980; 1983; Chang and Julesz, 1984; 1985; Williams et al., 1986; Williams and Phillips, 1987; Snowden and Braddick, 1989; Snowden 1989; Nawrot and Sekuler, 1989) and computational foundations (Grzywacz and Yuille, 1990).

Comparison with other species

Our results point to two stages of processing of motion information in monkey cortex. An extraction of the motion energy in each direction at each point in the image (performed in V1) followed by an interaction between directions of motion (evident in the response of MT neurons). Other species face similar demands in their environments and it may therefore be the case that a similar strategy for motion computation will be evident in other species. Of course the visual systems of other species vary immensely in their anatomy, so the site of these operations will be displaced relative to the monkey.

In the rabbit directionally selective cells can be found as early as the retinal ganglion cells (Barlow et al., 1964). Levick et al. (1969) performed an experiment similar in conception to those reported here on these cells. They found that the response to a spot of light moving in the cell's preferred direction was unaffected by the addition of a second spot moving in the null direction. However, if the same test was performed on the directionally selective cells of the rabbit LGN a strong inhibitory effect was found. Thus the response of the rabbit

retinal ganglion cells appears to mimic monkey V1 cells, and those of rabbit LGN mimic those of monkey MT.

The H1 neuron of the fly is believed to gain its directionally selective responses from the pooling of small field units termed elementary motion detectors (review: Franceschini et al., 1989) - a two stage process. The properties of each stage can be isolated by recording the response of the H1 neuron to wide field movement (the second, pooling stage) or to the same stimulus behind a narrow slit which eliminates pooling and thus isolates the elementary motion detectors (the first stage). Egelhaaf et al. (1990) measured the activity of both stages with and without the application of GABA agonists (GABA (g-aminobutyric acid) is a neurotransmitter thought to underlie inhibitory interactions: Bormann, 1988). They found that the response of the the first stage was unaffected by GABA agonists whereas the second, pooling stage was affected. Thus the first stage of the motion processing in the fly does not appear to use GABA inhibition, whereas the second stage does. This appears to a parallel our results of a lack of inhibitory effects at the first stage of motion processing in the monkey (V1) but inhibitory effects at the second stage (MT). Further, the action of GABA is thought to produce a shunting inhibition (Grzywacz and Koch, 1987; Schmid and Bülthoff, 1988) which would appear as a division-like operation in this area (see below).

In the cat directionally selective cells are first found in the striate cortex. Kaji and Kawabata (1985) have performed a similar experiment to ours on complex cells in cat striate cortex. While, no information is given concerning the population as a whole it is clear that some cells show a suppression of the response to their preferred direction when a stimulus of a different direction is superimposed. These cells

resemble our MT cells, and the V1 cells which showed a suppressive effect. The relative strength and abundance of the suppressive effect in cat and monkey striate cortex may well be different, with the monkey showing a much smaller effect (or fewer cells) than the cat. Gulyás et al., (1987) have shown that the response to a moving bar is suppressed by texture background patterns moving in the same direction in 55% of cat striate neurons, but a similar effect is observed in only 10% of monkey striate neurons (see also Orban et al., 1987). It therefore appears that some aspects of motion processing which occur in striate cortex of cats are delayed until MT in the monkey. Hence, the response to a bar can be suppressed by background texture in area MT (Tanaka et al., 1986), especially in the layers not receiving a direct projection from V1, i.e. outside layer IV (Lagae et al., 1989). Our results are consistent with these finding that many of the interactions required for this suppressive behaviour is performed for cells in area MT rather than those in area V1.

From these data it appears there are strong parallels in the way motion is processed between very different species, though the location of each stage of processing may be very different.

The nature of the inhibition

Inhibitory influences are widespread throughout cortex (Eccles,1969; Beneveto et al., 1972), and have already been shown to play a major role in shaping the response properties of neurons tuned to the orientation of a stimulus (Blakemore and Tobin, 1972; Sillito, 1977; 1979; Rose, 1977; Tsumoto et al., 1979; Burr et al., 1981; Morrone et al., 1982; Kaji et al., 1983; Ferster, 1986; Ramoa et al., 1986; Ferster and Koch, 1987; Bonds, 1989). One major question concerns the computational operation that inhibitory synapses perform. Blomfield (1974)

suggests that two basic operations can be performed depending upon the site of the inhibitory synapse (soma or dendrite) and their relationship to the site of excitatory innervation. One operation involves the linear addition of IPSPs and EPSPs and appears as a subtractive factor on the cells output, whereas the second operation involves shunting away the excitation flowing towards the soma and appears as a division factor (in the limiting cases). Our results (figures 5.12 & 5.14) show that the expected response to a single direction alone is reduced by addition of the antipreferred direction in such a manner that the expected output is divided by a constant factor (for a constant antipreferred stimulus) provided the response is not saturated. Increasing the strength of movement in the antipreferred direction increases this factor (figure 5.12). These results therefore favour the notion of a division-like inhibition (occurring before response saturation).

Division-like inhibition

The notion of a division-like action of inhibition is not new. Various cellular models have considered the neurotransmitter GABA (g-aminobutyric acid) to play such a role in the nervous system (Dreifuss et al., 1969; Krnjevi´c, 1974; Torre and Poggio, 1978; Grzywacz and Koch, 1987; Schmid and Bülthoff, 1988). Its action is thought to create low-resistance pathways through the cell membrane which acts to short-circuit some of the current flowing towards the soma. Thus recording intracellularly its effects would not be seen as IPSPs - hence it has often been referred to as silent or shunting inhibition (Torre and Poggio, 1978).

Division-like inhibition has been previously demonstrated in cat striate cortex by experiments which compare the response of neurons to single bars or gratings and when two such patterns

which differ in orientation are superimposed. It therefore appears that these interactions are important in the analysis of contours and spatial pattern (Morrone et al., 1982; Morrone et al., 1987; Ramoa et al., 1986). Rose (1977) has further shown that application of GABA agonists reduces orientation selectivity (see also Sillito, 1977; 1979) and that the reduction in response is proportional to the strength of the expected response (before application of GABA agonist) thus pointing to a division-like operation. Similar division-like interactions have also been suggested for interactions between oriented contours for human vision (Morrone and Burr, 1986; Burr and Morrone, 1987). Dean et al. (1980) also showed division-like inhibition between opposite directions of motion. They measured responses of cat area 17 neurons as a function of contrast for stimuli drifting in the preferred direction. They found that the addition of a second grating moving in the antipreferred direction made the function relating contrast to response to rise more gradually, but that the contrast threshold for a response did not change. This result is similar to our findings in area MT of the monkey.

Overall view of motion processing

Our results suggest that there is a motion stream through the primate cortex which is organized in a hierarchical fashion. Area V1 appears to extract motion energy in each direction at each point of the image, and then MT combines these estimates through inhibitory interactions. Thus V1 could provide the basis for segmenting the image, whereas the interactions in MT may serve to ease signal-to-noise problems (Snowden and Braddick, 1989) and smooth between motion estimates (Treue et al., 1991).

References

Adelson EH, Bergen JR (1985) Spatiotemporal energy models for the perception of motion. J. Opt. Soc. Am. A 2: 284-299.

Adelson EH, Movshon JA (1982) Phenomenal coherence of moving visual patterns. Nature. 300: 523-525.

Albright TD (1984) Direction and orientation selectivity of neurons in visual area MT of the macaque. J. Neurophysiol. 52: 1106-1130.

Allman J, Miezin F, Mc Guiness E (1985a) Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). Perception. 14: 105-126.

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

Amthor FR, Gryzwacz NM (1991) Nonlinearity of the inhibition underlying retinal directional selectivity. Vis. Neurosci. 6: 197-206.

Andersen RA, Snowden RJ, Treue S, Graziano M (1990a) Hierarchical processing of motion in the visual cortex of monkey. Cold Spring Harbor Symposia on Quantitative Biology, LV: 1990.

Andersen, RA, Asanuma CC, Essick G, Siegel RM (1990b) Corticocortical connections of anatomically defined subdivisions within the inferior parietal lobe. J. Comp. Neurol. 296: 65-113.

Andersen RA, Bracewell RM, Barash S, Gnadt JW, Fogassi L (1990c) Eye position effects on visual, memory, and saccade-related activity in areas LIP and 7a of Macaque. J. Neurosci. 10: 1176-1196.

Andersen RA, Siegel, RM (1990) Perception of structure-from-motion in monkey and man. J. Cog. Neurosci. 2: 306-319.

Ando H, Hildreth EC, Treue S, Andersen RA (1990) Recovering 3-D structure from motion with surface reconstruction. Soc. Neurosci. Abstr. 16: 962.

Baker CL Jnr, Braddick OJ (1985) Eccentricity-dependant scaling of the limits for short range apparent motion. Vision Res. 25: 803-812.

Baker CL Jnr, Cynader MS (1988) Space-time seperability of direction selectivity in cat striate cortex neurons. Vision Res. 28: 239-246.

Baker CL Jnr, Hess RF, Zihl J (1990) The "motion-blind" patient: Perception of random dot "limited lifetime" motion. Invest. Ophthalmol. Vis. Sci. Suppl. 31: 239.

Barlow HB, Levick WR (1965) The mechanism of directionally selective units in the rabbit's retina. J. Physiol. (Lond.). 178: 477-504.

Barlow HB, Hill RM, Levick WR (1964) Retinal Ganglion cells responding selectively to direction and speed of image motion in the rabbit. J. Physiol. (Lond.). 173: 377-407.

Benevento LA, Creutzfeldt OD, Kuhnt V (1972) Significance of intracortical inhibition in the visual cortex. Nature. 238: 124-126.

Blakemore C, Tobin EA (1972) Lateral inhibition between orientation detectors in the cat's visual cortex. Exp. Brain Res. 15: 439-440.

Blomfield S (1974) Arithmetic operations performed by nerve cells. Brain Res. 69: 115-124.

Bonds AB (1989) Role of inhibition in the specification of orientation selectivity of cells in the cat striate cortex. Visual Neurosci. 2: 41-55.

Bormann J (1988) Electrophysiology of GABAa and GABAb receptor subtypes. Trends Neurosci. 11: 112-116.

Borst A, Egelhaaf M (1989) Principles of visual motion detection. Trends Neurosci. 12: 297-306.

Braddick OJ (1974) A short-range process in apparent motion. Vision Res. 14: 519-527.

Britten KH, Newsome WT (1990) Responses of MT neurons to discontinous motion. Invest. Ophthalmol. Vis. Sci. Suppl. 31: 238.

Bülthoff H, Little J, Poggio T (1989) A parallel algorithm for real-time computation of optical flow. Nature. 337: 549-553.

Burr DC, Morrone C, Maffei L (1981) Intra-cortical inhibition prevents simple cells from responding to textured visual patterns. Exp. Brain Res. 43: 455-458.

Burr DC, Morrone MC (1987) Inhibitory interactions in the human visual system revealed with pattern-evoked potentials. J. Physiol. (Lond.). 389: 1-21.

Chang JJ, Julesz B (1984) Cooperative phenomena in apparent motion perception of random-dot cinematograms. Vision Res. 24: 1781-1788.

Chang JJ, Julesz B (1985) Cooperative and non-cooperative processes of apparent motion of random dot cinematograms. Spatial Vision. 1: 39-45.

Clarke PGH (1977) Subjective standstill caused by interaction of moving patterns. Vision Res. 17: 1243.

De Valois RL, Yund EW, Kepler NK (1982) The orientation and direction selectivity of cells in macaque visual cortex. Vision Res. 22: 531-544.

Dean AF, Hess RF, Tolhurst DJ (1980) Divisive inhibition involved in direction selectivity. J. Physiol. (Lond.). 308: 84P.

Douglas RJ, Martin KAC, Whitteridge D (1988) Selective responses of visual cortical cells do not depend on shunting inhibition. Nature. 332: 642-644.

Dow BM, Vautin RG, Bauer R (1985) The mapping of visual space onto foveal striate cortex in the macaque monkey. J. Neurosci. 5: 890-902.

Dreifuss JJ, Kelly JS, Krnjevi´c K (1969) Cortical inhibition and g-aminobutyric acid. Exp. Brain Res. 9: 137-154.

Eccles JC (1969) The inhibitory pathways of the central nervous system.Liverpool: Liverpool UP.

Egelhaaf M, Borst A, Pilz B (1990) The role of GABA in detecting visual motion. Brain Res. 509: 156-160.

Emerson RC, Citron MC, Vaughin WJ, Klein SA (1987) Nonlinear directionally selective subunits in complex cells of cat striate cortex. J. Neurophysiol. 58: 33-65.

Erickson RG, Snowden RJ, Andersen RA, Treue S (1989) Directional neurons in awake rhesus monkeys: Implications for motion transparency. Soc. Neurosci. Abstr. 15: 323.

Fennema CL, Thompson WB (1979) Velocity determination in scenes containing several moving objects. Comp. Graph. Image Proc. 9: 301-315.

Ferster D (1986) Orientation selectivity of synaptic potentials in neurones of cat primary visual cortex. J. Neurosci. 6: 1284-1301.

Ferster D, Koch C (1987) Neuronal connections underlying orientation selectivity in cat visual cortex. Trends Neurosci. 10: 487-491.

Franceschini, N., Riehle, A. & Nestour, A. L. (1989) Directionally selective motion detection by insect neurons. In: Facets of vision, Stavenga,and Hardie, eds. pp 360-390, Spinger-Verlag, Heidelberg.

Fries W, Keizer K, Kuypers HGJ (1985) Large layer V1 cells in macaque striate cortex (Meynert cells) project to both superior colliculus and prestriate visual area V5. Exp. Brain Res. 58: 613-616.

Gallyas F (1979) Silver staining of myelin by means of physical development. Neurol. Res. 1: 203-209.

Ganz L, Felder R (1984) Mechanism of directional selectivity in simple cells of the cat's visual cortex analyzed with stationary flash sequences. J. Neurophysiol. 51: 294-324.

Gattass R, Gross CG (1981) Visual topography of striate projection zone (MT) in posterior temporal sulcus of the macaque. J. Neurophysiol. 46: 621-638.

Golomb, B., Andersen, R. A., Nakayama, K., MacLeod, D. I. A. & Wong, A. (1985) Visual tresholds for shearing motion in monkey and man. Vision Res. 25: 813-820.

Goodwin AW, Henry GH, Bishop PO (1975) Direction selectivity of simple striate cells: Properties and mechanism. J. Neurophysiol. 38: 1500-1523.

Goodwin AW, Henry GH (1975) Direction selectivity of complex cells in a comparison with simple cells. J. Neurophysiol. 38: 1524-1540.

Grimson WEL (1981) From images to surfaces. A computational study of the human early visual system. MIT press. Cambridge, MA.

Grzywacz NM, Amthor FR (1989) Facilitation in on-off directionally selective ganglion cells of the rabbit retina. Soc. Neurosci. Abstr. 15: 969.

Grzywacz NM, Koch C (1987) Functional properties of models for direction selectivity in the retina. Synapse. 1: 417-434.

Grzywacz NM, Yuille AL (1990) A model for the estimate of local image velocity by cells in the visual cortex. Proc. Roy. Soc. Lond. B 239: 129-161.

Gulyás B, Orban GA, Spileers W (1987) A moving noise background modulates responses of striate neurons to moving bars in the cat but not the monkey. J. Physiol. (Lond.). 390: 28P.

Hammond P, MacKay DM (1981) Modulatory influences of moving textured backgrounds on responsiveness of simple cells in feline striate cortex. J. Physiol. (Lond.). 319: 431-442.

Hammond P, Smith AT (1983) Directional tuning interactions between moving oriented and textured stimuli in complex cells of feline striate cortex. J. Physiol. (Lond.). 342: 35-49.

Hawken MJ, Parker AJ, Lund JS (1988) Laminar organization and contrast sensitivity of direction-selective cells in the striate cortex of the Old World monkey. J. Neurosci. 8: 3541-3548.

Heeger DJ (1987) Model for the extraction of image flow. J. Opt. Soc. Am. A. 4: 1455-1471.

Hildreth EC (1984) The computation of the velocity field. Proc. Roy. Soc. Lond. B 221: 189-220.

Hoffmann K-, Distler C (1989) Quantitative analysis of visual receptive fields of neurons in nucleus of the optic tract and dorsal terminal nucleus of the accessory tract in macaque monkey. J. Neurophysiol. 62:416-428.

Horn BKP, Schunk BG (1981) Determining optical flow. Artificial Intelligence. 17: 185-203.

Hubel D, Wiesel T (1968) Receptive fields and functional architecture of the monkey striate cortex. J. Physiol. (Lond.). 195: 215-243.

Husain, M., Treue, S. & Andersen, R. A. (1989) Surface interpolation in three-dimensional structure-from-motion perception. Neural Comp. 1: 324-333.

Hutchinson J, Koch C, Luo J, Mead C (1988) Computing motion using analog and binary resistive networks. IEEE Computer. 21: 52-61.

Judge SJ, Richmond BJ, Shu FC (1980) Implantation of magnetic search coils for measurement of eye position: An improved method. Vision Res. 20: 535-537.

Kaji S, Kawabata N (1985) Neural interactions of two moving patterns in the direction and orientation domain in the complex cells of cat's visual cortex. Vision Res. 25: 749-753.

Kaji S, Tamane S, Kawabata N (1983) Neural interactions of two slits in the orientation domain in the visual units of the cat. Vision Res. 23: 883-886.

Koch C, Poggio T, Torre V (1986) Computations in the vertebrate retina: gain enhancement, differentiation and motion discrimination. Trends Neurosci. 9: 204-211.

Kowler E, van der Steen J, Tamminga EP, Collewijn H (1984) Voluntary selection of the target for smooth eye movement in the presence of superimposed, full-field stationary and moving stimuli. Vision Res. 24:1789-1798.

Krnjevi´c K (1974) Chemical nature of synaptic transmission in vertebrates. Physiol. Rev. 54: 418-540.

Lagae L, Gulyás B, Raiguel S, Orban GA (1989) Laminar analysis of motion information processing in macaque V5. Brain Res. 496: 361-367.

Lappin JS, Kottas BL (1981) The perceptual coherence of moving visual patterns. Acta Psychol. 48: 163-174.

Levick WR, Oyster CW, Takahashi E. (1969) Rabbit lateral geniculate nucleus: sharpener of directional information. Science, 165: 712-714

Levinson E, Sekuler R (1976) Adaptation alters perceived direction of motion. Vision Res. 16: 779-781.

MacKay DM (1961) Visual effects of non-redundant stimulation. Nature. 192: 739-740.

Marshak WM, Sekuler R (1979) Mutual repulsion between moving visual targets. Science. 205: 1399-1401.

Mather G (1980) The movement aftereffect and a distibution-shift model for coding the direction of visual movement. Peception. 9: 379-382.

Mather G, Moulden B (1980) A simultaneous shift in apparent directions: Further evidencve for a 'distribution-shift' model of direction coding. Quart. J. Exp. Psychol. 32: 325-333.

Mather G, Moulden B (1983) Thresholds for movement direction : Two directions are less detectable than one. Quart. J. Exp. Psychol. 35: 513-518.

Maunsell JHR, Van Essen DC (1983) The connections of the middle temporal visual area (MT) and their relationship to a cortical hierarchy in the macaque monkey. J. Neurosci. 3: 2563-2586.

Mikami A, Newsome WT, Wurtz RH (1986a) Motion selectivity in macaque visual cortex. I. Mechanisms of direction and speed selectivity in extrastriate area MT. J. Neurophysiol. 55: 1308-1327.

Mikami A, Newsome WT, Wurtz RH (1986b) Motion selectivity in macaque visual cortex: II. Spatio-temporal range of directional interactions in MT and V1. J. Neurophysiol. 55: 1328-1339.

Morgan MJ, Ward R (1980) Interocular delay produces depth in subjectively moving noise patterns. Quart. J. Exp. Psychol. 32: 387-395.

Morrone MC, Burr DC, Maffei L (1982) Functional implications of cross-orientation inhibition of cortical cells. I. Neurophysiological evidence. Proc. Roy.Soc. Lond. B 216: 335-354.

Morrone MC, Burr DC, Speed HD (1987) Cross-orientation inhibition in cat is GABA mediated. Exp. Brain Res. 67: 635-644.

Morrone MC, Burr DC (1986) Evidence for the existence and development of visual inhibition in humans. Nature. 321: 235-237.

Movshon JA, Adelson EH, Gizzi MS, Newsome WT (1985) The analysis of moving visual patterns.In: Pattern recognition Mechanisms, C. Chagas, R. Gattass and C. Gross eds., Rome: Vatican Press. Pont. Acad. Sci. Scr. Varia. 54: 117-151.

Movshon JA, Thompson ID, Tolhurst DJ (1978) Receptive field organization of complex cells in the cat's striate cortex. J. Physiol. (Lond.). 283: 79-99.

Nawrot M, Sekuler R (1989) Assimilation and contrast in motion perception: explorations in cooperativity. Invest. Ophthalmol. Vis. Sci. Suppl. 30: 72.

Newsome WT, Britten KH, Movshon JA (1989) Neuronal correlates of a perceptual decision. Nature. 341: 52-54.

Newsome WT, Mikami A, Wurtz RH (1986) Motion selectivity in macaque visual cortex. III. Psychophysics and physiology of apparent motion. J. Neurophysiol. 55: 1340-1351.

Newsome WT, Paré EB (1988) A selective impairment of motion perception following lesions of the middle temporal visual area (MT). J. Neurosci. 8: 2201-2211.

Orban GA, Gulyás B, Spileers W (1987) A moving noise background modulates responses to moving bars of monkey V2 cells but not of monkey V1 cells. Invest. Ophthalmol. Vis. Sci. Suppl. 28: 197.

Pasternak T, Albano JE, Harvitt DM (1990) The role of directionally selective neurons in the perception of global motion. J. Neurosci. 10: 3079-3086.

Poggio T, Koch C (1987) Synapses that compute motion. Scientific American.256: 46-52.

Ramoa AS, Shadlen M, Skottun BC, Freeman RD (1986) A comparison of inhibition in orientation and spatial frequency selectivity of cat visual cortex. Nature. 321: 237-239.

Reichardt W (1961) Autocorrelation: a principle of the evaluation of sensory information by the central nervous system. In: Sensory Communication. (Rosenblith WA ed.) New York: Wiley.

Robinson DA (1963) A method of measuring eye movement using a sceral search coil in a magnetic field. IEEE Trans. Biomed. Eng. 10: 137-145.

Rodman HR, Albright TD (1989) Single-unit analysis of pattern-motion selective properties in the middle temporal visual area (MT). Exp. Brain Res. 75: 53-64.

Rose D (1977) On the arithmetic operation performed by inhibitory synapses onto the neuronal soma. Exp. Brain Res. 28: 221-223.

Saito, H. -A, Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y. & Iwai, E. (1986) Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. J. Neurosci., 6: 145-157.

Schmid A, Bülthoff H (1988) Using neuropharmacology to distinguish between excitatory and inhibitory movement detection mechanisms in the fly Calliphora erythrocephala. Biol. Cybern. 59: 71-80.

Shipp S, Zeki S (1985) Segregated output to area V5 from layer 4B of macaque monkey striate cortex. J. Physiol. (Lond.). 369: 32P.

Siegel RM, Andersen RA (1986) Motion perceptual deficits following ibotenic acid lesions of the middle temporal area in the behaving rhesus monkey. Soc. Neurosci. Abstr. 12: 1183.

Siegel RM, Andersen RA (1988) Perception of three-dimensional structure from motion in monkey and man. Nature. 331: 259-261.

Sillito AM (1977) Inhibitory processes underlying the directional specificity of simple, complex and hypercomplex cells in the cat's visual cortex. J. Physiol. (Lond.). 271: 699-720.

Sillito AM (1979) Pharmacological approach to the visual cortex. Trends Neurosci. 2: 196-198.

Snowden RJ (1989) Motions in orthogonal directions are mutually suppressive. J. Opt. Soc. Am. 7: 1096-1101.

Snowden RJ, Braddick OJ (1989) The combination of motion signals over time. Vision Res. 29:1621-1630.

Snowden RJ (1990) Suppressive interactions between moving patterns: Role of velocity. Percept. Psychophys. 47: 74-78.

Snowden RJ, Braddick OJ (1990) Differences in the processing of short-range apparent motion at small and large displacements. Vision Res. 30: 1211-1222

Snowden RJ, Erickson RG, Treue S, Andersen RA (1990) Transparent stimuli reaveal divisive inhibition in area MT of macaque. Invest. Ophthalmol. Vis. Sci. Suppl. 31: 399.

Stromeyer CF III, Kronauer RE, Madsen JC, Klein SA (1984) Opponent-movement mechanisms in human vision. J. Opt. Soc. Am. A 1: 876-884.

Sutherland NS (1961) Figural after-effects and apparent size. Quart. J. Exp. Psychol. 13: 222-228.

Tanaka K, Hikosaka K, Saito H, Yukie M, Fukada Y, Iwai E (1986) Analysis of local and wide-field movements in the superior temporal visual areas of the macaque monkey. J. Neurosci. 6: 134-144.

Torre V, Poggio T (1978) A synaptic mechanism possibly underlying directional selectivity to motion. Proc. Roy. Soc. Lond. B 202: 409-416.

Treue S, Husain M, Andersen R (1991) Human perception of structure from motion. Vision Res. 31: 59-75.

Tsumoto T, Eckart W, Creutzfeldt OD (1979) Modification of orientation sensitivity of cat visual cortex neurons by removal of GABA-mediated inhibition. Exp. Brain Res. 34: 351-363.

Uras S, Girosi F, Verri A, Torre V (1988) A computational approach to motion perception. Biol. Cybern.60: 79-87.

Vaina LM, LeMay M, Bienfang DC, Choi AY, Nakyama K (1990) Intact "biological motion" and "structure from motion" perception in a patient with impaired motion mechansims: A case study. Visual Neurosci. 5: 353-369.

van Doorn AJ, Koenderink JJ, van der Grind WA (1985) Perception of movement and correlation in stoboscopically presented noise patterns. Perception. 14: 209-224.

Wang HT, Mathur B, Koch C (1989) Computing optical flow in the primate visual system. Neural Computation. 1: 92-103.

Watson AB, Ahumada AJ Jnr (1985) Model of human visual-motion sensing. J. Opt. Soc. Am. A. 1: 322-342.

Watson AB, Thompson PG, Murphy BJ, Nachmias J (1980) Summation and discrimination of gratings moving in opposite directions. Vision Res. 20: 341-347.

Williams D, Phillips G, Sekuler R (1986) Hysteresis in the perception of motion direction as avidence for neural cooperativity. Nature. 324: 253-255.

Williams P, Phillips P (1987) Cooperative phenomena in the perception of motion direction. J. Opt. Soc. Am. 4: 878-885.

Wohlgemuth A (1911) On the aftereffect of seen motion. Brit. J. Psychol. 1: 1-117.

Yuille AL, Grzywacz NM (1988) A computational theory for the perception of coherent visual motion. Nature. 333: 71-74.

Yuille AL, Grzywacz NM (1989) A winner-take-all mechanism based on presynaptic inhibition feedback. Neural Computation. 1: 334-347.

Zeki SM (1974) Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. J. Physiol. (Lond.). 236: 549-573.

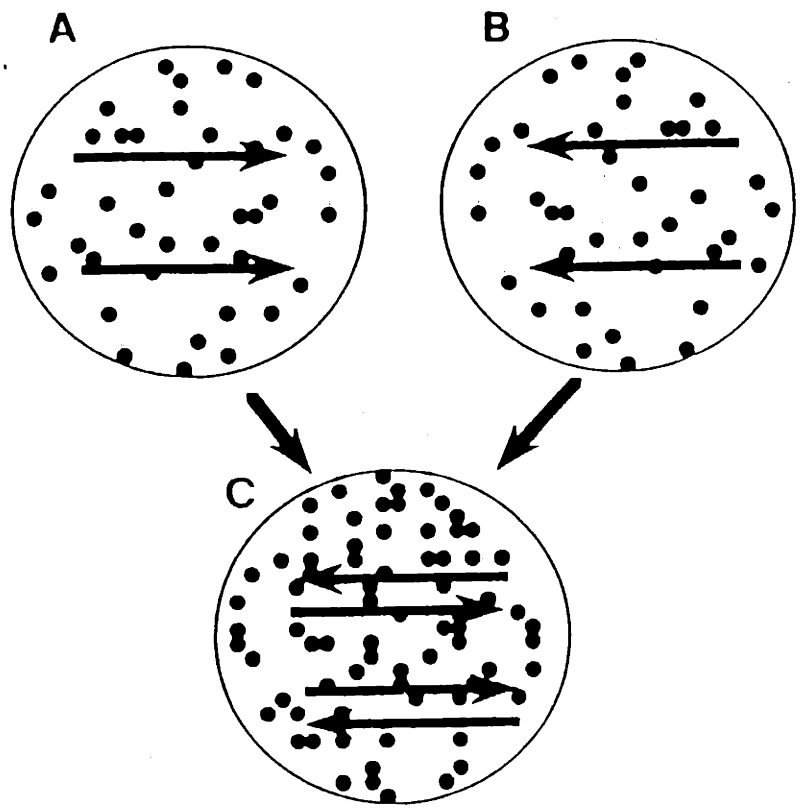


Figure 5.1

A cartoon representation of the stimuli used in this study. Dots were randomly placed within a window and shifted every frame by a set amount. Those falling outside the original window were wrapped to the opposite side of the display. For the single surface stimuli (A & B) all dots were displaced by the same amount and in the same direction. The two-surface stimuli (e.g. C) had two sets of dots which could undergo separate manipulations. In the stimulus portrayed in c the dots moved in the same speed in opposite directions, and thus this stimulus is equivalent to the superimposition of the stimuli portrayed in A & B. In other experiments the dots could be made to move at any arbitrary angle to one another, to have different densities in each surface, to be stationary or randomly repositioned every other frame or to be positioned into discrete sections of the pattern.

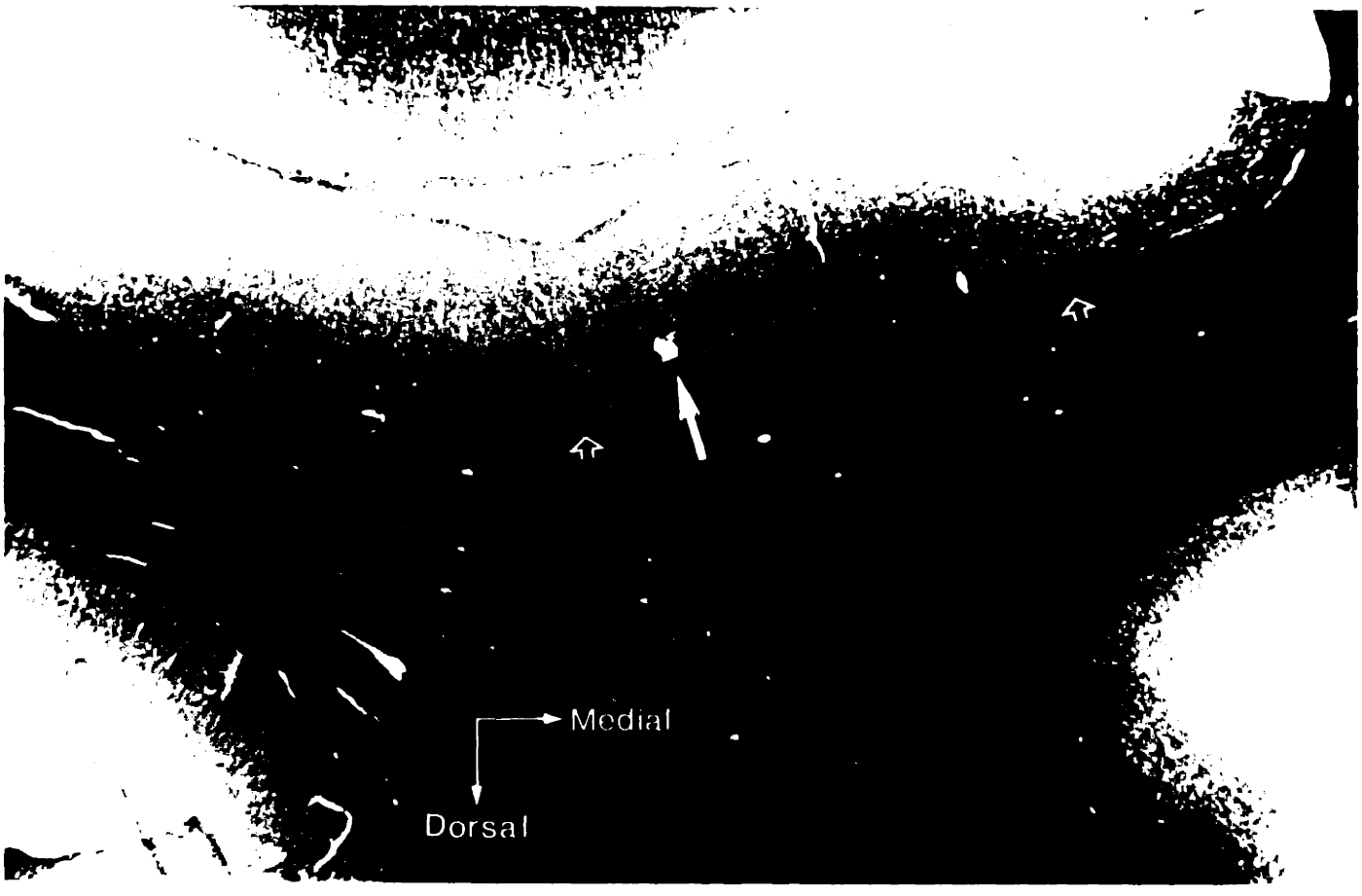


Figure 5.2

Horizontal section through the superior temporal sulcus. Area MT is visible by its distinctive heavy myelination on the posterior bank of the sulcus (marked by the hollow arrows in the underlying white manner). A marking lesion is visible in the lateral half of MT (filled arrow).

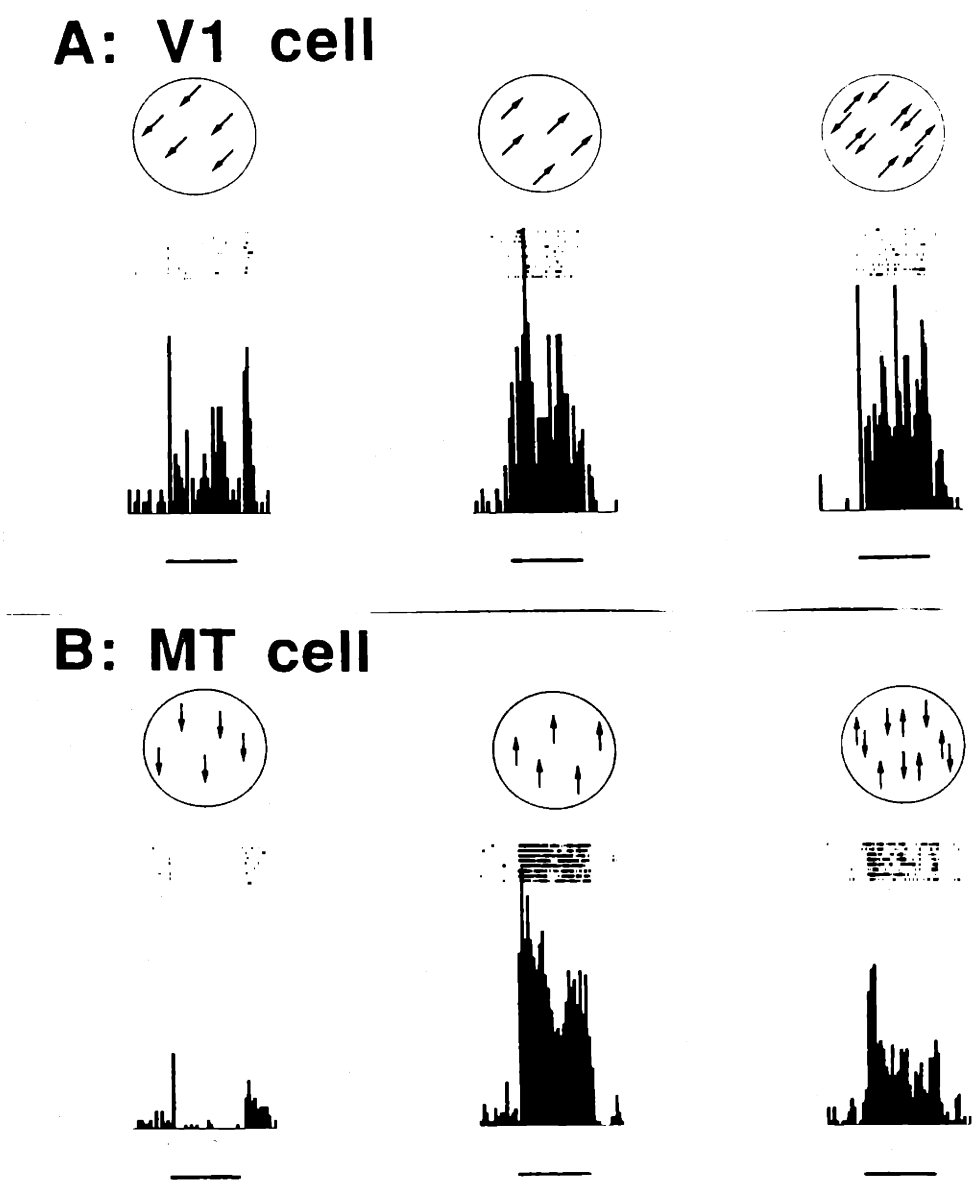


Figure 5.3

Examples of the responses of a V1 and a MT neuron to single and two surface stimuli. The circles containing arrows depict the type of stimulus shown on those trials and the dark bars below the response histograms indicate the time the stimulus was displayed. Each major division of the abscissa is 1000 msec, and each division of the ordinate is 100 spikes per sec. For the V1 cell (A) the cell gives a directional response (favouring movement up and to the right) and gives a response to the two-surface stimulus which is similar in magnitude to the preferred stimulus alone. Thus this cell had an Id of 0.82 and an Is of 0.16 (see text for details). The MT cell (B) is also directional (Id =1.03) but gives a lower response to the two-surface stimulus than it does to the preferred stimulus alone (Is = 0.55).

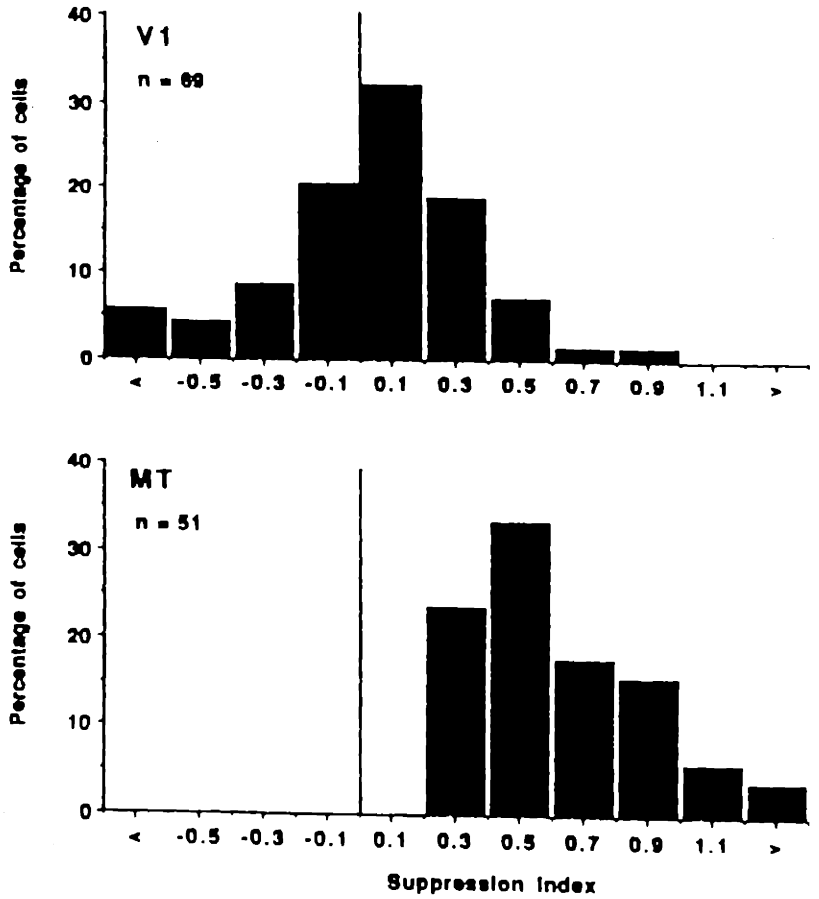


Figure 5.4

For each of our cells on which the suppression test was performed we calculated the Is, and the frequency of occurrence of this index is plotted for our population of V1 cells (upper section) and MT cells (lower section). A Is of 0.0 (indicated by the vertical line) means that the cell responded in a similar fashion to the preferred stimulus alone and the two-surface stimulus. Points to the left of this line indicate a greater response to the two-surface stimulus, whilst points to the right indicate a smaller response to the two-surface stimulus. The median index for V1 was 0.04 and for MT 0.54. The distribution of the Is in our population of V1 and MT cells is significantly different ($p < .0001$).

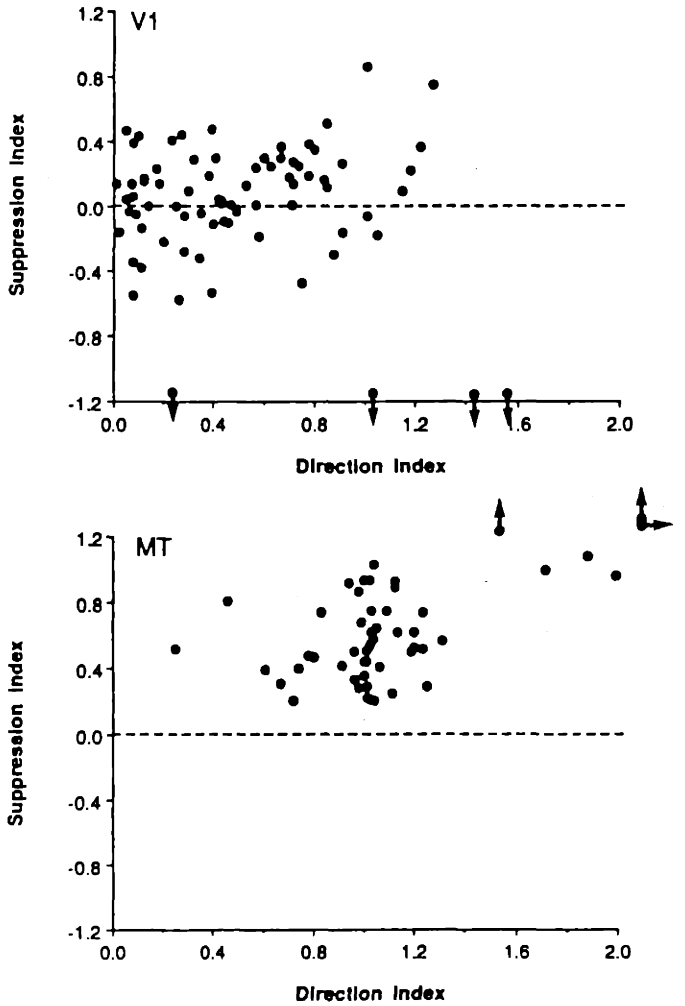


Figure 5.5

A scatter plot of the Id against the Is for the V1 cells (upper section) and MT cells (lower section). Points which fell of the scale are indicated by arrows. A statistical test for correlations (Spearman's rank) was performed between the two indexes. The V1 data show no significant correlation between the two indexes ($p > 0.5$) whereas the MT cells do ($p < 0.005$) which indicates that cells with a large Id also tend to have a large Is.

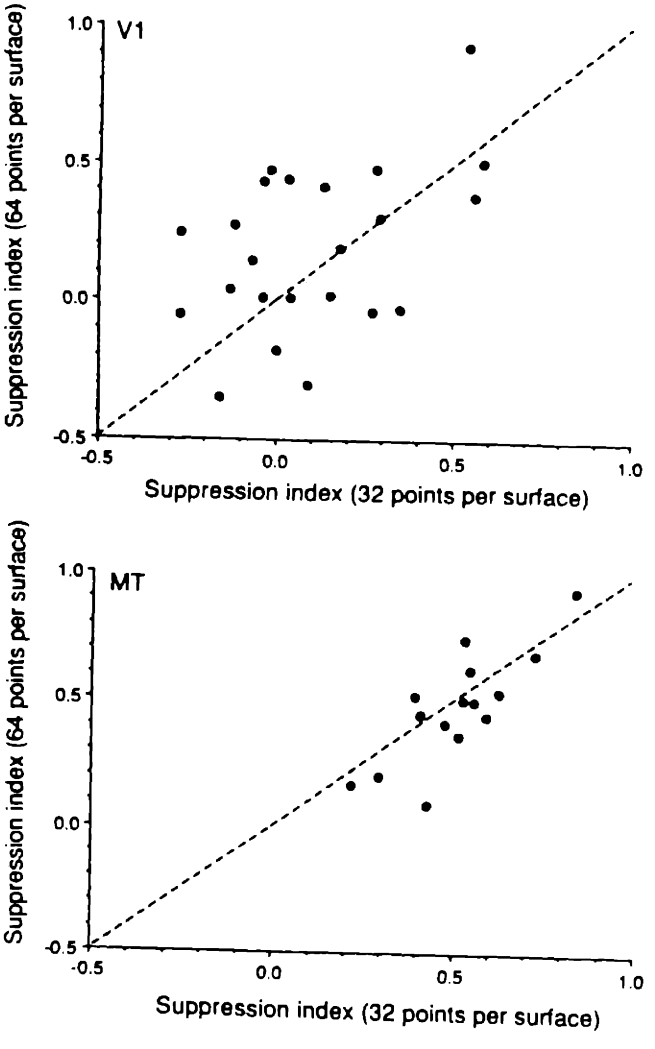


Figure 5.6

The Is calculated when the surfaces in the trasparent condition contained 64 points each is plotted against the same index when the surfaces contained 32 points each. The single-surface stimulus used to determine the response to the preferred direction alone contained 64 points in both cases. The upper section contains the results from V1 cells and the lower section MT cells. Both plots show a strong correlation between the two indexes and points fall with approximately equal preference above and below the dotted line at 45° which indicates no systematic shift in the Is with dot density. The V1 data have a greater scatter than that of the MT data and this scatter was found to increase with decreasing dot density. This scatter may reflect the statistical probabilities of the number of dots in the receptive fields of the V1 neurons.

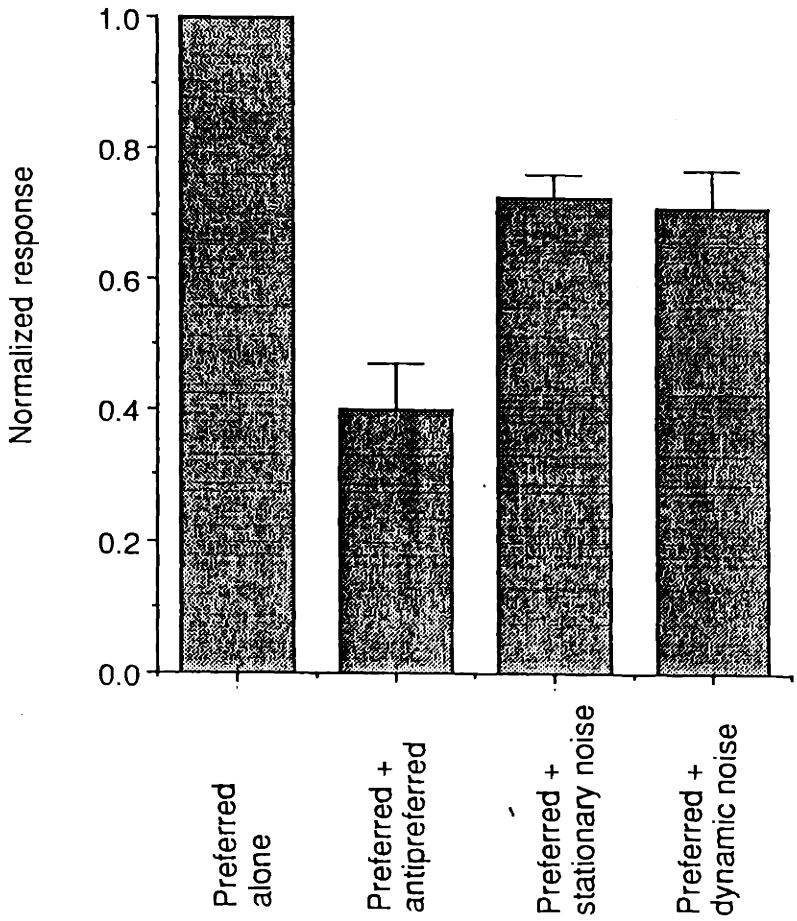


Figure 5.7

A comparison of the effects of adding the antipreferred direction of movement, stationary noise and dynamic noise upon the response of MT cells (N = 13) to its preferred direction of movement. All stimuli were presented in a pseudorandom order. The response elicited in condition was scaled to the response of the preferred stimulus alone (the preferred stimulus alone therefore has a score of 1.0). The mean index is plotted in the form of a histogram and the error bars indicate one standard error. A statistical analysis (Wilcoxon signed-rank) has shown significant differences between the preferred alone and all three other conditions, between the antipreferred and the stationary and dynamic, but no significant differences between the stationary and dynamic noise conditions.

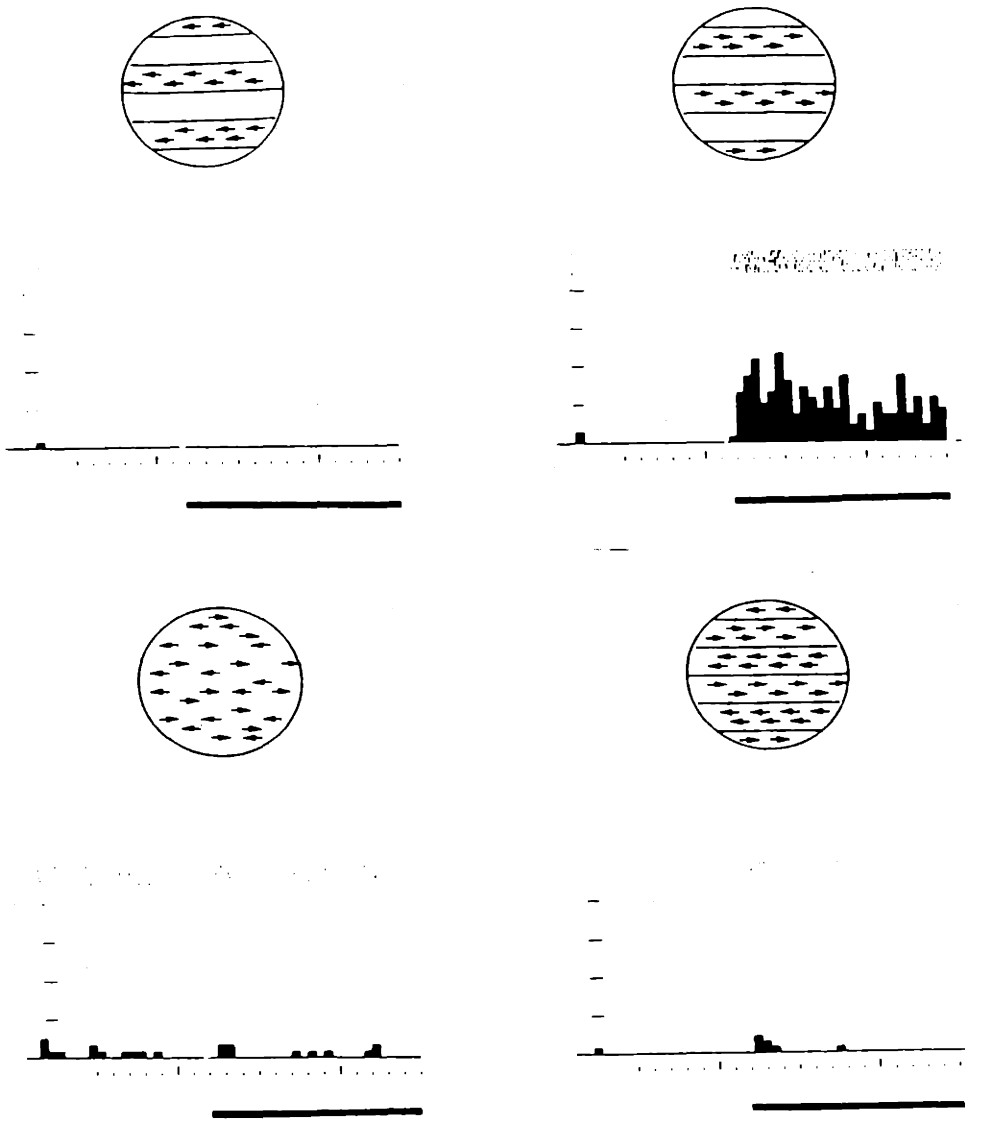


Figure 5.8

An example of the responses of an MT cell to motion which is kept in discrete stripes. In the upper sections only a single direction of motion is present in a stimulus, though the dots are contained within three discrete stripes. Data were also collected when the stripes which are blank in this diagram were the ones filled with dots, and this data was identical in form to that portrayed here. The lower sections of the figure show the response of the cell when two motions were within the receptive field (all conditions were pseudorandomly interleaved). Both when the motion vectors were overlapping and when they were in discrete sections we found a strong suppressive effect.

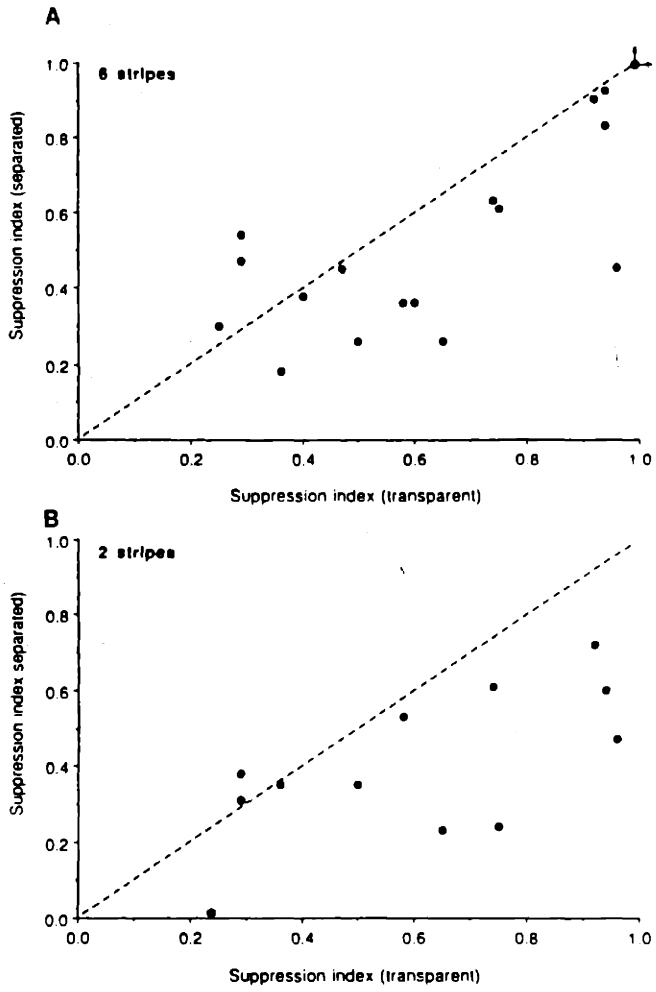


Figure 5.9

The Is calculated when the motions were superimposed (transparent) and when the motions were in discrete stripes (separated) are plotted against one another for MT cells. If the indexes are exactly the same points should lie along the dotted line at 45°. If the suppression effect were to disappear under the separated condition the points should lie along the abscissa.

A) Here the dots were separated into 6 stripes (3 moving in each direction. Clearly the points lie much closer to the 45° line indicating a strong suppression effect even under the discrete condition. There is, however, a small trend towards having a smaller Is under the discrete condition.

B) Here the dots are separated in to 2 stripes (1 moving in each direction). The results are similar to those of a.

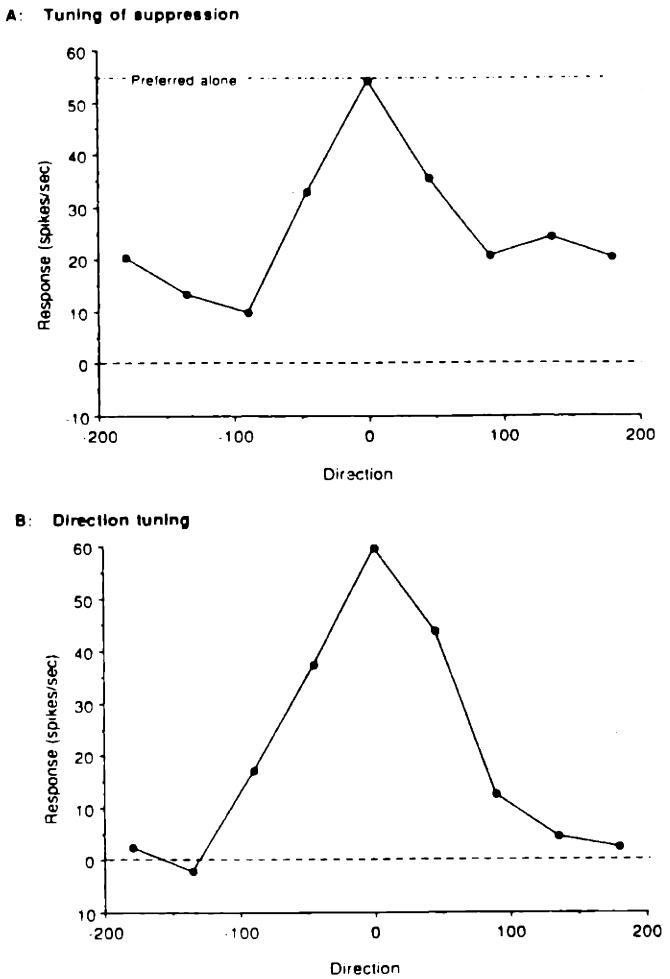


Figure 5.10

A) The directional tuning of the suppression effect for a MT cell. The upper dot-dash line indicates the response of the cell to the preferred direction alone and the points indicate the response of the cell when each of 8 directions were superimposed upon this preferred motion (the point for motion 180° from preferred is plotted twice). As can be seen all motions other than the preferred itself reduced the response of the cell as compared to the preferred direction alone indicating that the data so far presented is probably not unique to opposite directions of motion. For this cell the largest effects were actually found for motions of orthogonal direction.

B) As a comparison the direction tuning of the same cell as in a is presented. The direction tuning was produced by simply drifting a single surface pattern in each of 8 directions (again the 180° point is repeated). The slight differences in peak response are probably due to the fact that these two curves were gathered at separate times (not in a pseudo-interleaved paradigm as other results). It is clear that motions at 45° and 90° to the preferred direction produce a strong excitatory response. These are exactly the same motions that produce an inhibitory effect in part A.

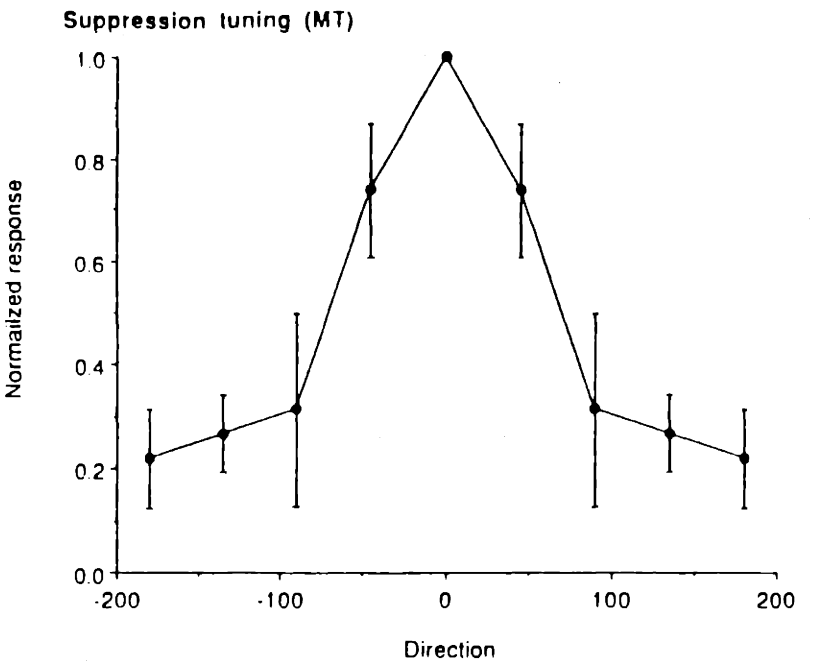


Figure 5.11

The direction tuning of the suppressive effect across a population of MT neurons (N = 11). The test for each neuron was that described in figure 10a. The response in each condition was then scaled to the response to the preferred direction, and the mean and standard error are portrayed in the figure.

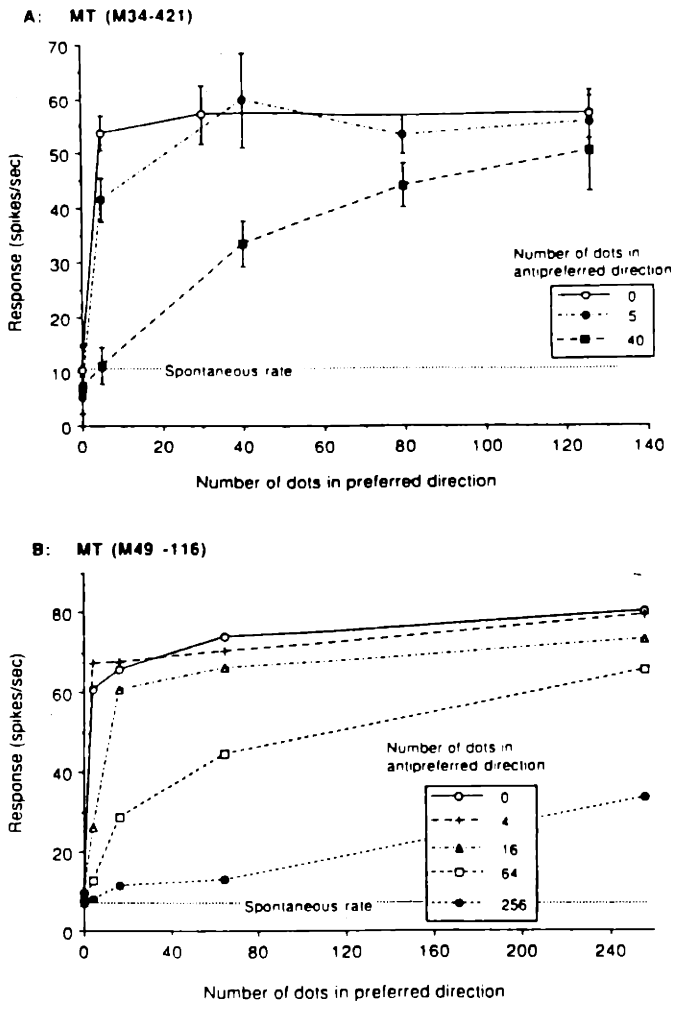


Figure 5.12

The response of two MT cells (A & B) is plotted as a function of the number of dots in the preferred direction. Each curve corresponds to repeating this function with a certain number of dots always drifting in the antipreferred direction (the number is given on the figure). All trials for each cell were presented in a pseudorandom order. The standard deviation of each point is plotted in section A but omitted in section B in order to avoid cluttering. When no dots are drifted in the antipreferred direction the curves rise sharply and asymptote with increasing dot density of the preferred direction. The addition of dots in the antipreferred direction causes this rise to be more gradual but does not necessarily reduce the maximum response of the cell (though this is now reached at a greater dot density). Increasing the dot density in the antipreferred direction increases these effects.

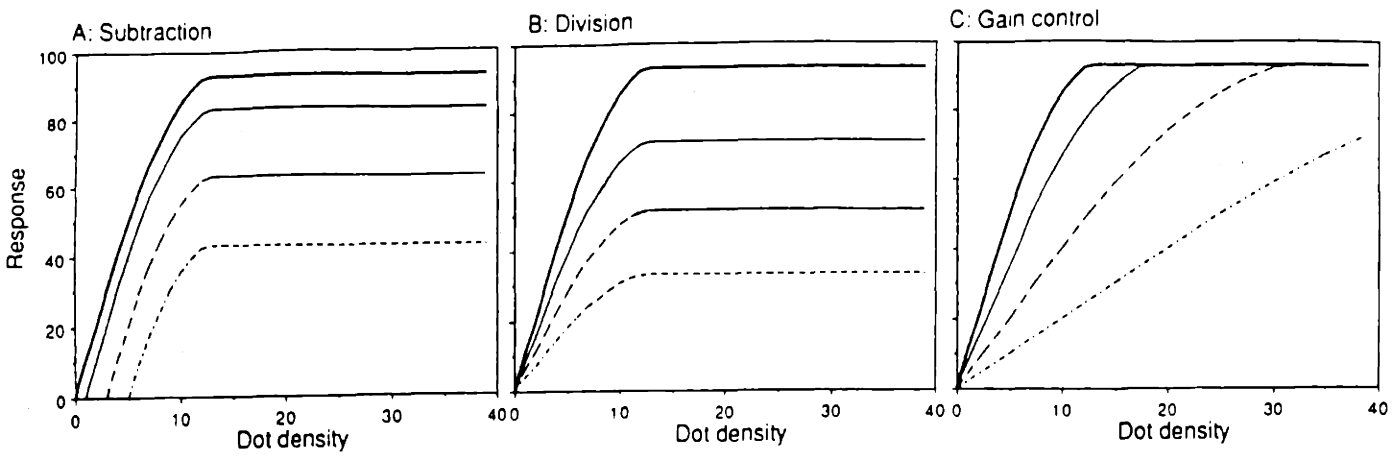


Figure 5.13

A simulation of the expected effects of various forms of inhibition upon the response versus dot number curves for cells.

A) The curve (thick line) simulates the function produced by the preferred direction alone. It rises quickly with increasing dot density and then begins to saturate and finally asymptotes. This approximates the action of real MT neurons (see figure 12) and is similar in form to data obtained form cat and monkey V1 neurons for response versus contrast (Albrecht and Hamilton, 1982). This curve is also reproduced in sections B and C. The other curves represent the actions of a purely subtractive inhibition, with curves further down the vertical axis representing greater inhibition.

B) As for section A only the curves now represent the action of a division type operation occurring after the mechanism of saturation.

C) As for A only the curves represent the action of a division-type operation occurring before the mechanism of saturation. Such a process can be considered as a gain control.

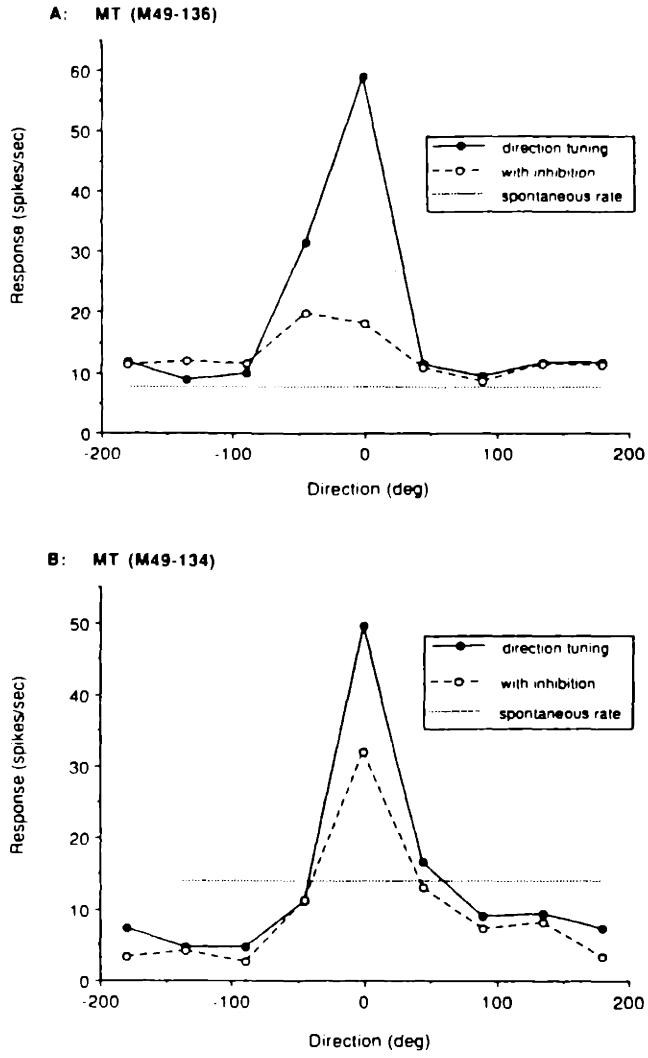


Figure 5.14

The direction tuning of two MT cells under conditions of just a single surface (solid line and symbols), and where an equal number of dots were always drifting in the antipreferred direction (180°, open symbols dashed line). the spontaneous rate is also shown as a dotted line. The greatest suppression occurs at 0° when the cell would be most active.

Chapter 6

The Response of Neurons in Areas V1 and MT of the Alert Monkey to Moving Random Dot Patterns

Introduction

Complex patterns such as visual noise and random dot patterns have been widely used in the psychophysical research of motion perception because they contain no features which can be tracked from frame to frame, thus allowing the isolation of the motion system (Nakayama and Tyler 1981). Their use for single cell recording has also proved fruitful (e. g. Hammond and Mackay 1975; Gulyás et al. 1987; Skottun et al. 1988) but has been limited to the anesthetized cat. While these studies are most valuable, it is often difficult to compare such data with psychophysical studies of motion perception in humans. Two major factors seem relevant. Firstly, when a psychophysical task is performed the subjects typically attempt to fixate a point. The exact point of fixation will vary from trial to trial (as we will demonstrate for the monkey), and there will be microsaccades and slow drifts even when 'fixation' occurs (Motter and Poggio 1984; Snodderly and Kurtz 1985). Secondly, striate cortex in the cat and monkey cannot be regarded as functionally identical. For example the proportion of directionally selective cells is different in the two areas (around 70% in the cat and 30% in the monkey and in addition, they have different layering characteristics (Gilbert 1977; Hawken et al. 1987). It therefore is of great interest to collect data on the response of single cells in the alert primate and to compare these data with relevant studies of human psychophysics, the physiology of cat cortex, and with previous reports on monkey cortex using both anesthetized and alert animals. One notable exception to the studies cited above is a recent report by Newsome et al. (1989). They recorded from neurons in MT/V5, an area believed to be important for the perception of motion (Zeki 1974; Newsome and Paré 1988) while a monkey viewed moving, partially

correlated random dot patterns. They were able to show that individual neurons can, with certain assumptions, discriminate opposite directions of motion with a similar degree of accuracy as the animal itself.

In order to ascertain the relationships between psychophysical performance and the response properties of single cells we have recorded from single units in areas MT and V1 of the alert, behaving monkey. The stimuli employed were patterns of randomly placed dots of high luminance, whose direction, density and speed were under experimental control. We have also quantified the variance of eye movements while the animals fixated in the presence of the stimuli. The response variance of area V1 and MT cells was measured for different mean levels of activity and the direction tuning of area MT neurons was quantified.

Using the data from single MT cells relating the mean response to the direction of motion, and the response variance to the mean response, we modelled the ability of individual MT neurons to discriminate different directions of motion. Direction discrimination thresholds for human observers were also measured using the same stimuli employed in the single cell studies. To anticipate the results, our data indicate that a small number of MT neurons can discriminate changes in direction with a similar precision to human observers. Interestingly, the minimum discriminanda occur on the flanks rather than on the peaks of their curves. Since Newsome et al. (1989) showed that detection of motion direction is mediated by cells with peaks in their tuning curves in that direction, these experiments suggest that different populations of MT neurons are responsible for threshold detection and suprathreshold discrimination of motion.

Material and Methods

Overview

A detailed description of our recording methods has appeared elsewhere (Snowden et al., 1991) and this section will therefore be limited to a brief overview and a more detailed description of the stimuli used.

Two male rhesus monkeys were trained to fixate a small fixation point, while ignoring the test motion stimuli, and to signal the dimming of the fixation point by releasing a key. Using a scleral search coil technique (Robinson 1963) the animals' eye movements and point of fixation were closely monitored. Visual stimulation was provided to the receptive field of individual neurons during this 4 - 6 sec period of fixation. Electrode penetrations were made through a chamber implanted over area V1. This placement allowed us to sample cells from V1, V2 and MT. Over the course of many penetrations topographic maps of each area were compiled and these were used as an aid to assigning each recording site to an area. During the final 6 penetrations of one monkey marking lesions (all occurring within the 2 weeks prior to the animal being sacrificed) were placed at relevant sites and these were used to help reconstruct recording sites after histological reconstruction.

Stimuli and data analysis

Stimuli consisted of bright dots (30 ft. lambert) randomly plotted upon a dark background. Each dot was approximately 1 mm in diameter, and subtended 6 min arc. The pattern was circular and subtended 3 deg at the viewing distance of 57 cm. Under most conditions to be reported a total of 64 dots were

used which corresponds to a dot density of 7% or 9.2 dots/deg. This type of pattern is similar to that used by Skottun et al. (1988) but somewhat different from that used by some previous investigators (e.g. Hammond and MacKay 1975; 1977) where the texture was made by assigning each pixel black or white (50% dot density).

Movement was created by displacing the X and Y coordinate of each element by a certain amount. Dots which would have fallen outside the 3 deg circle were wrapped to the opposite side of the display. Each element had a limited point lifetime of 500 msec, after which it was randomly replotted on some other part of the screen. The rate of screen refresh was 60 or 35 Hz. Each trial commenced with the onset of the fixation point. After 1 sec the stimulus appeared if the animal was successfully fixating. This stimulus was extinguished after 1 sec, and another stimulus appeared for 1 sec after a 1 sec delay. The fixation point dimmed 0.2 - 2.0 secs after the end of the last stimulus; thus a complete trial lasted 4.2 - 6 secs. In this manner we were able to present two stimuli per trial. This was the case for most of our data; however, for the earliest recordings (consisting of approximately 20 % of V1, and 40 % of the MT recordings) a single stimulus was presented for 3-5 secs.

The response to the stimulus was calculated for a 1 sec period whose commencement was aligned with response onset. For each stimulus condition 6 - 10 trials were completed and the mean response and standard deviation were calculated. Only cells whose activity could be significantly modulated by random dot patterns are included in this study. To test for a significant response modulation we performed a t-test between two conditions (A and B) via the formula:

$$A_{mean} - B_{mean} \; / \; ((A_{SD}^2 / N_A) + (B_{SD}^2 / N_B))^{1/2}$$

where sd is the standard deviation and Na and Nb the number of trials run on conditions a and b respectively. Any cell that scored > 5.0 on any such test was included in the study.

Psychophysics

In addition to the neurophysiological studies described above we also collected human psychophysical data using the same stimuli so that more direct comparisons could be made than would be afforded by using data from other laboratories. A single random dot pattern was presented on each trial. On half the trials the pattern moved 'upward' and on half 'downward' in order to prevent aftereffects. The pattern did not move vertically but with an angle slightly to the left or right of the vertical, and the subject's task was to indicate which of these had been presented via a button box press (binary forced choice). No feedback on performance was given. In order to help the subject maintain a good sense of the vertical, a stationary dot was placed 0.5 deg above and below the pattern along the vertical axis. Nine subjects took part in the experiment. Seven subjects were given a very brief (2- 4 min) practice session to familiarize themselves with the task, and two of the subjects were well practiced. The method of constant stimuli was employed. Twenty stimuli (ten up, ten down) were presented 20 times each in pseudorandom order and the proportion judged moving to the right of vertical were plotted as a function of angle (figure 6.9b; 90° = vertical) separately for the upward and downward conditions. The data were fitted by an integrated Weibull function and the discrimination threshold was taken as the angle change from a probability of 0.5 to one of 0.75. The two discriminanda were averaged for upward and downward motion for each subject.

Results

Eye position

In the introduction we suggested that the alert animal has differences in eye movement over the paralysed animal. To quantify this for our monkeys we recorded the point of eye position fixation from trial to trial over a series of 100 successive trials. The eye position at the time of the fixation point dimming was taken as the measure of the point of fixation. Table 1 presents the standard deviations for both vertical and horizontal position of the eye for each monkey. Both animal's showed variation in fixation of only a few min of arc in good accord with previous measurements (Motter and Poggio 1984; Snodderly and Kurtz 1985). It should be noted that some of the scatter found in the position might be attributable to the inherent noise in the measuring system. We attempted to estimate this by collecting eye position data using a coil, similar to the one implanted in the monkey's eye, placed in the magnetic field. The standard deviation in the output was around 1 DAC (digital to analogue conversion) unit, which corresponds to 0.80 mins in the vertical axis and 0.87 mins of monkey 34, and to 1.3 mins in both axes of monkey 49. Thus the standard deviation measured for the monkey's fixation contained a small component due to noise in the measuring system.

Table 1

Standard deviations of eye positions from trial to trial [min arc]

Monkey number	Horizontal eye position	Vertical eye position
34	5.9	6.3
49	4.3	8.2

While the trial to trial variability in eye position is quite small this represents a substantial fraction of receptive field size for area V1 neurons. At the eccentricities we recorded from in V1 (1 - 3 deg) receptive fields vary in size from about 10 mins to around 1 deg (Dow et al 1981).

Response properties

1) Grain of response

Nearly all the cells we formally tested were driven to some extent by random dot patterns. It is hard; however, to know the percentage of cells which were unresponsive to such stimuli as much of our searching and mapping of receptive fields was performed using these patterns. Hence, there is a systematic bias towards finding this type of cell.

While nearly all V1 cells tested were driven by the random dot patterns it was clear that different types of response could occur. Many cells did not respond in a uniform manner during the stimulus presentation time (see figure 6.1a). Instead these cells tended to fire at a certain time (and not at other times) during stimulus presentation. As the same pattern was presented on each trial this is consistent with the cell firing to some particular feature or phase relationship within the dot

pattern. Such a result has been previously reported for cat V1 cells (Hammond and MacKay 1975; Gulyás et al. 1987) where this type of response was termed a 'grain' response. Other V1 cells (e.g. figure 6.1b) gave responses which were much more consistent over the time course of the stimulus. Gulyás et al. (1987) term this a 'field' response. Nearly all MT cells gave responses that were similar to this "field" response. Quantifying the 'graininess' of the response is not a trivial operation (as the response is very much tied to the exact pattern used) and so in order to give some impression of the relative amounts of graininess in each area we rated (3 (very grainy) - 0 (no grain)) the graininess of each cell (the observer did not know from which area the cell came). The result is displayed in figure 6.1c and shows that very grainy response types were confined to area V1.

2) Effects of dot density

One possible explanation of our finding of a greater grain type of response in area V1 is that, since receptive fields are much smaller in area V1 as compared to MT (Gattass and Gross 1981), our stimuli often extended beyond the boundaries of the V1 receptive fields. This leads to a smaller mean number of dots in the receptive field which would increase the variability of how many dots actually were within the receptive field at any given time. At the extreme end would be a receptive field so small that it would either contain one dot or no dot at all. Such effects might account for the greater incidence of grain type responses in V1 cells than in MT cells. We investigated this issue by systematically recording the response of cells as a function of dot density of the pattern. We found that cells that gave a grain type response did so even with a sixteen-fold increase in dot density and we therefore suggest that it is not the scarcity of dots which induces this graininess (clearly if very low dot

densities are used then a graininess can be induced in any cell). It was also noticeable in most cells that changes in dot density seemed to make little difference to firing rates. To quantify this we normalized the response of each cell to its maximum firing rate and calculated the mean response and the standard deviation at each dot density tested in each area. Figure 6.2a and b plots these curves for areas V1 and MT respectively. Both functions show a rapid rise with increasing dot density and a saturation at a fairly low dot density. There are no obviously differences between areas V1 and MT. It should be further noted that our normal number of dots (64) lies in the area which produces a saturated response.

3) Direction tuning

In order to compare our results with previous estimates of the directionality of cells in the cortex it is necessary to derive an index which can be used across studies. Unfortunately a standard index has not yet been adopted by all laboratories and this complicates our attempts to compare indices derived with random dot patterns with those obtained using other stimuli.

The direction index (Id) we chose to use was:

$$Id = 1 - A / P \quad (1)$$

where P stands for the firing rate in the preferred direction and A the firing rate in the opposite (antipreferred) direction. This index was calculated after the spontaneous rate (obtained when the animal was fixating an otherwise blank screen) had been subtracted. Values near 0.0 indicate no difference between these directions (a non-directional cell), and increasing values indicate greater and greater directionality. With this index it is also possible to quantify when motion in the antipreferred direction causes the cell to be suppressed below the spontaneous rate (values > 1.0).

Figure 6.3 plots the occurrence of Id in both V1 and MT. As has been noted by many other authors (Zeki 1974; Albright 1984; Mikami et al. 1986) the cells of MT show a far greater directionality than those of V1 (median V1 = 0.44; median MT = 1.01; Mann-Whitney U: $p < 0.0001$).

We have considered a cell to be directional if it gave a response which was three times greater for the preferred direction than for the null direction. About 32% of the V1 cells, and 93% of MT cells, gave such a response. These figures are in excellent agreement with previous studies on the monkey (Schiller et al. 1976; De Valois et al. 1982a; Hawken et al. 1988) using bar and grating stimuli, suggesting that such estimates are relatively independent of the type of stimulus used.

De Valois et al. (1982a) presented data suggesting a bimodal distribution of directionality within the macaque V1 with many cells showing just a weak preference for direction and a few cells showing a strong preference with essentially no response to the anti-preferred direction. Our data exhibit no sign of this bimodality. The reason for this discrepancy may lie in the different stimuli employed in this study (dot patterns) and their studies (lines and gratings); however, other factors such as anesthesia may also be involved. The direction index used by De Valois et al. (antipreferred/preferred response) is very sensitive to small changes in firing rate for low responses in the antipreferred direction. The use of anesthetics has been previously shown to reduce the activity of cells (Livingstone and Hubel 1981). This reduction could serve to increase the number of cells with extreme direction indices, generating their bimodal distribution. It should be noted that the data of Hawken et al. (1988), who performed experiments very similar to those of De Valois et al. (1982a) but used another directionality measure,

show no bimodality.

Variance

Neurons of the visual cortex fire in a probabilistic fashion. That is to say that identical stimuli do not produce exactly the same response from trial to trial. This variability of the response to a stimulus (and in the cells' spontaneous rate) has been of tremendous interest (Heggelund and Albus 1978; Tolhurst et al. 1981, 1983; Dean 1981; Parker and Hawken 1985; Bradley et al. 1987; Scobey and Gabor 1989; Vogels et al. 1989; Zohary et al. 1990) because it helps determine and constrain the capacity of a cell to signal the presence or absence of features/information in the world.

The majority of the studies cited above recorded the response of cells in striate cortex of anesthetized cats. The effect of anesthetic on response properties is still not totally clear (e. g. Livingstone and Hubel 1981) but must serve to complicate any comparison between performance measures of single neurons, especially when attempts are made to compare this performance with those of the behaviour of an animal. A further complication is that paralysis of the eye muscles is induced. Small eye movements still occur under conditions of strict fixation and there is a limited ability to fixate precisely the same point from trial to trial (see table 1; Motter and Poggio 1984). This introduces variance into the precise nature of the position of the stimulus with respect to the receptive field of the neuron being examined which is not present for the paralysed preparation, but may be a factor in limiting psychophysical thresholds. Secondly, the functional organization of visual cortex is sufficiently different in cats and primates that we are most wary of comparing these data to human psychophysical

performance. Taken together this suggests that the optimal data to compare with human psychophysics are those recorded from an alert, fixating primate performing some psychophysical task.

The first report of such efforts has recently appeared. Vogels et al. (1989) report upon the relationship between response and variability for a population of V1 neurons stimulated with stationary square wave gratings. In this section we extend these data by examining the response to moving random dot patterns. We have also examined the response of MT cells and establish (for the first time) the variability of responses in an extrastriate area.

For neurons where recording was stable for a long enough period we have examined the relationship between variance and mean response. Figure 6.4 shows the variance (the square of the standard deviation) as a function of mean response on a double logarithmic plot for both representative V1 and MT neurons. The variance increases with increasing mean response for both cells. In accord with previous studies (Tolhurst et al. 1981; Dean 1981; Vogels et al. 1989) we attempted to fit a power function of the form:

variance = x * response y (2)

where y represents the slope of the straight line on log-log coordinates and x the intercept (i.e. the variance when mean response = 1). In both areas we found such a power law to be an adequate fit to the data, and clearly superior to other relationships (i. e. semilog, linear, exponential).

Such plots were made for all suitable neurons. Figure 6.5 shows frequency histograms of the power functions (A and C) and intercepts (B and D) encountered in both areas. In area V1 we obtained a mean power of 1.21 (sd = 0.28, N = 41) and intercept of 1.08 (sd = 0.77, N = 41).

The mean slope in area MT was 1.10 (sd = 0.29, N = 39) and the mean intercept 1.37 (sd = 1.55, N = 39). This is the first estimate of such a parameter for extrastriate visual cortex and is highly suggestive that the variance to response relationship in extrastriate cortex is essentially the same as that found in striate cortex. A non-parametric test (Mann- Whitney U) failed to find any significance in the difference between the data from V1 and MT for the slopes ($p = 0.24$) or the intercepts ($p = 0.43$) and therefore we have no grounds for believing that the variance in extrastriate area MT is dissimilar to striate cortex.

Each point in the variance-versus-response functions displayed in figure 6.5 was produced by obtaining responses to a certain random dot pattern. In order to manipulate changes in the response level we varied such factors as the direction of movement, speed of movement and dot density. One question of interest is therefore if changes along different stimulus dimensions produce different functions. If they do then clearly we are not justified in pooling across the various stimulus dimensions. Figure 6.6 demonstrates variance versus response functions from an MT cell produced by changes along the dimensions of speed (open symbols), and direction of motion (closed symbols) separately. In this instance it is clear that both dimensions produced very similar functions. Similar tests were applied to 13 MT neurons and a non-parametric test (Mann-Whitney U) revealed no significant difference between either the slope ($p > 0.5$) or intercept ($p > 0.5$) when produced by changing the stimulus along these two dimensions. A similar result was found by Dean (1981) who showed that changes along the dimensions of spatial frequency and contrast produce similar functions in cat striate neurons (see also Vogels et al. 1989).

Further, the similarity between the slopes produced in the present study using random dot patterns, and those of other studies cited above using gratings, all suggest that it is the mean response level that determines the variance and the value of the variance has little to do with how that mean response level is produced.

Direction discrimination by MT cells

In the previous section we addressed the variability of response rate for both V1 and MT cells. The most notable physiological characteristic of MT cells is their different response rates to different directions of motion. Thus these cells have the potential to discriminate the direction of motion of a pattern, and may well play a vital role in determining psychophysical thresholds for the monkey including that of direction discrimination. The ability of a cell to discriminate directions of motion is dependent upon how much its firing rate changes with changes in stimulus direction, and upon the reliability of its response. In theory, if a cell gave exactly the same response to identical stimuli then discrimination would be limited only by the quantum nature of spike generation. However, as we demonstrated in the previous section, this is far from the case and the variance of a cell is often of the same order as the mean response. In this section we therefore attempt to estimate the capacity of MT neurons to discriminate direction of motion by producing neurometric functions (Tolhurst et al. 1981).

Neurometric functions can be estimated in at least two ways. The first is via the receiver operating characteristic (ROC) curve (e. g. Bradley et al. 1987). Typically two stimuli are presented and the probability that the cell gives a greater response to one of the stimuli is observed. From this the cell's ability to

discriminate small changes can be calculated. This method requires a very large number of trials and is not particularly suited to work on the alert animal. The second method is to measure the underlying characteristics of the cell, such as its tuning along a particular dimension and its response variability, and then use these measurements to model the response of the cell to various stimuli (e.g. Scobey and Gabor 1989). The latter method requires many fewer trials and is therefore more suited to experiments on the alert animal. We therefore chose this method.

Direction tuning curves were determined by measuring the response of each neuron at 8 different directions of motion for 6 - 10 trials per direction (see methods) and plotting the mean response as a function of direction (e.g. figure 6.7). The data were then fitted by a Gaussian function of the form:

$$\text{response} = r_{min} + r_{max} * \exp\left(-0.5 * d^2 / s^2\right) \qquad (3)$$

where d is the angle of motion away from the preferred direction (degrees) , r_{min} is the minimum firing rate, r_{max} the maximum firing rate, and s the standard deviation of the Gaussian. We found this function to be an excellent fit to all but two of our sample of neurons (N = 32). The cells which were poorly fit appeared to be so due to a significant bi-directionality in their response, and we eliminated them from further analysis. The finding that a Gaussian provides an excellent fit to MT direction tuning curves is in agreement with Albright (1984) who sampled at 16 different directions of motion. For our population of cells the mean standard deviation was 46.5 deg (sd = 18.2, N = 30) which is slightly larger than that obtained by Albright (who reported a full bandwidth at half height of 85 deg, which is equivalent to a sd of 36.3 deg).

Knowing the relationship between direction of motion and mean response (3), and the function relating mean response to variance (2), we were in a position to simulate the response of an MT cell to a stimulus moving in any direction. We ran simulations in which 36 directions of motion were presented. For each stimulus a mean response was calculated according to the parameters obtained from equation 2, and then a number was picked from a Gaussian distribution[1] (mean = 0, sd =1 unit) and scaled according to equation 1. These two numbers were then added to produce the response on that trial. In total 20,000 trials were simulated for each cell. From these data neurometric functions could be produced (e.g. figure 6.8a+b). For a certain criteria (e.g. number of spikes on a trial) we calculated the probability that more than this number of spikes was elicited for each direction. These data points were fitted by the formula:

$$P = g - (g - d) * \exp\left(-1 * (d/a)^b \right) \qquad (4)$$

where d is the direction of motion, a the direction at which a criterion probability is reached, b the parameter governing the slope of the function, d the asymptotic value of P (i.e. when d = 0), and g the probability of reaching criterion for the antipreferred direction (i. e. d = 180). This equation is the integral of the Weibull function. These neurometric functions are equivalent to psychometric functions in that they describe the cell's ability to respond differentially to different directions of motion. Thus the range of directions over which the cell can change the probability of a criterion response by 25 percent is equivalent to a direction discrimination threshold. We chose the

[1] The distribution of responses to a particular stimulus is not a perfect Gaussian (Bradley et al. 1987; Dean 1981; Scobey and Gabor 1989). However, the deviation from a Gaussian distribution is not great at high firing rates, the Gaussian distribution is mathematically convenient, and it has been used successfully in similar modelling attempts (Scobey and Gabor 1989).

range f directions that changes the probability from 0.5 to 0.25 since that covers the steepest portion of the curve.

As can be seen in figure 6.8a and 6.8b such a function can be generated for any arbitrary criterion (provided the probability of response passes through 0.5 and 0.75). Each of these functions yields a direction discrimination threshold and these are plotted in figure 6.8c as a function of the criterion for the cells illustrated in figure 6.8a and b. Discrimination thresholds follow a U-shaped function with a broad base over which discrimination is finest and almost constant. The position of this shallow minimum is for criteria which fall at some distance away from the peak firing rate, and therefore the preferred direction, of the cell. This is because near the peak of the Gaussian function the slope is reduced and, as firing is near maximum, the variance is the greatest.

We used the region of the U-shaped threshold criterion plot (Fig. 6.8b) to estimate the discrimination ability for each cell[2]. Since we are determining a cell's ability to discriminate different directions of movement, rather than simply detect the presence of movement, we used stimuli that evoked large responses when they moved in the preferred direction of the cell. This is desirable since the shape of the response distribution deviates markedly from a Gaussian for low response values (Dean 1981; Scobey and Gabor 1989; Tolhurst et al. 1983; Bradley et al. 1987). Thus as responses are high, and we are comparing stimuli which are producing similar response levels (and therefore response distributions) this issue is diffused. Clearly, a similar analysis on a detection task, like the estimation

2 This approach is supported by psychophysical evidence which suggests that some discrimination thresholds are determined by cells which respond maximally to stimuli other than the ones to be determined (Regan and Beverley 1983).

of contrast thresholds would be less valid as response rates would be low. In such a situation one would need to employ signal detection theory.

Figure 6.9a plots the frequency of occurrence of direction discrimination thresholds for our population of MT cells. In addition we measured direction discrimination thresholds for nine human observers (see methods) using the same stimuli as were employed during the recording sessions. Figure 6.9b plots a psychometric function for an individual observer. The data points have been fitted by the same integrated Weibull function as was employed for the neurophysiological data. As can be seen, the thresholds are approximately 1 deg. As a population the mean discrimination threshold was 1.1 (SD = 0.4; N = 9) which is similar to the figures provided by de Bruyn and Orban (1988) and Ball et al. (1983) when corrected to the same criterion. Comparison of this figure (indicated by the arrow in figure 6.9a) with the results portrayed in figure 6.9a shows that while there is considerable scatter in the thresholds of individual neurons there appear to be cells whose discrimination abilities are of the order of that of human observers.

Discussion

The nature of the response to random patterns in V1.

Cells of area V1 show a range of response types to drifting random dot patterns. Many cells respond in an intermittent manner showing several irregular bursts of firing over the time course of the stimulus. One possibility is that the response of these cells is not a true 'texture' response, but rather a response to a particular feature or phase relationship within the pattern. Such a response can only be identified if the same stimulus is used from trial to trial; therefore those studies which refresh the noise pattern between trials would miss this. In addition the position of the eyes must also be stable from trial to trial and could be missed due to slow drifts in eye position in the paralysed animal. Our ability to show this grain-type response testifies to the ability of our monkeys to fixate the same point in a consistent manner from trial to trial. Whether a cell responds with a grain-like response or with a field-like response may depend upon the acuity of the cell and the fineness of the texture employed (Hammond and Pomfrett 1989). While we found no obvious changes in graininess by a sixteen fold increase in dot density (1.75 - 28%), this is still different from the 'visual noise' employed by Hammond and colleagues (e.g. Hammond and MacKay 1975) and by Orban and colleagues (e.g. Gulyás et al. 1987) where each pixel is assigned black or white (i.e. 50% dot density). However, Gulyás et al. (1987) using such a pattern (pixel size = 2.4 mins) still classified cells into 3 classes depending upon their response to texture, no response (22%), grain response (55%), and field response (22%). So it appears grain responses can still exist at high dot densities. Our results from monkey striate cortex are in broad agreement with this study since we find cells which failed to respond, those giving a

grainy response, and those giving a sustained response (figure 6.1).

Field-type responses could occur from the neural convergence of signals from many grain-type cells in a manner similar to that suggested for how the phase invariant complex cell could be produced by the averaging of many phase variant simple cells (Holub and Morton-Gibson 1981; De Valois et al. 1982b). Such a hierarchical process from phase-dependent to phase-invariant responses has been argued to be of great importance in motion processing (Borst and Egelhaaf 1989) and is consistent with several recent models of human motion perception (van Santen and Sperling 1985; Adelson and Bergen 1985). For example the motion-energy model of Adelson and Bergen (1985) has several stages, with each stage predicting a particular type of response to random dot patterns. At the first stage of the model (separable responses) the response to random dot patterns is phase sensitive and non-directional. The response to random dot patterns is therefore grainy and non-directional for a time averaged response (though at any particular moment in time it could appear directional). The next stage (oriented linear response) gives a greater response to motion in a particular direction (for a time averaged response). However, as this stage is linear it is still phase sensitive (gives a grainy response to random dot patterns) and has two undesirable effects: 1) it has opposite preferred directions of motion for stimuli of opposite contrast (e. g. Albus 1980): and 2) at any particular instance it is hard to determine directionality due to the grain type response to random dot patterns. The next stage of the model (oriented energy) is produced in the model by summing the squared output of two directional filters whose phase preference is shifted by 90°. This stage is phase independent and gives a constant output throughout time. The

final stage of the Adelson and Bergen model (the opponent energy stage) consists of differencing the output from 'oriented energy' cells of opposite preferred direction. Interactions between different directions of motion have previously been demonstrated (Snowden et al 1991) and are much more prevalent in area MT than in area V1.

Dot density

Figure 6.2 shows that dot density has a relatively minor role in changing response strength. In both area MT and V1 responses rise quickly with dot density and saturate at low dot densities (though there is considerable variability between neurons). This is in accord with psychophysical results which show that dot density has very little effect in determining the upper and lower displacement limits of apparent motion (Baker and Braddick 1982) or in determining signal to noise ratios for detecting motion (Downing and Movshon 1988).

Variance

Our results (figs. 6.4-6) show that the variance of a cell to a particular stimulus is proportional to (and just a little greater than equal to) the mean response. The average slope of the function was 1.21 for V1 cells and is in very good agreement with previous studies of the anesthetized cat (Dean 1981; Tolhurst et al. 1981; 1983; Scobey and Gabor 1989) and of the alert monkey (Vogels et al. 1989). However, our estimation of the average intercept parameter (1.08) is around half those previously cited (above references). This may arise from using alert animals. Many previous studies (e. g. Tolhurst et al. 1983) have noted that the observed response variance may be an overestimation due to slow changes in the responsiveness of the cells over time.

MT cells are thought to receive an excitatory drive from many V1 cells. If many of these V1 cells are driven by a stimulus, and their response/variance characteristics are as described above, then it should be possible for each MT cell to derive a much lower variance to mean response as the noise associated with each V1 cell should cancel while the signal should add (assuming the noise is uncorrelated from cell to cell). However, our estimates of the variance to mean response characteristics of MT cells is very similar to that of V1 cells; that is the variance is nearly proportional to the mean response. Hence our results show no sign of improvement due to pooling. This result suggests that noise associated with the response of a MT cell may arise from mechanisms inherent in the cell itself, rather than being inherited from its inputs. Similar response variance ratios have been found elsewhere in cortex (Werner and Mountcastle 1963).

Direction discrimination

By using the fits to the functions relating mean response to direction, and relating variance to mean response, we were able to simulate the response of MT cells to different motion directions. The analysis shows that the cells can discriminate directions of motion which are but a small fraction of the tuning bandwidths. A similar result has been found for orientation discrimination and orientation bandwidths (Parker and Hawken 1985; Bradley et al. 1987; Scobey and Gabor 1989). While the above analysis is suggestive of the information processing capacity of MT neurons it should not be taken as a fait accompli. There are certainly some problems in applying these estimates directly to psychophysics. The first is that the analysis is based upon counting spikes over a time period of 1 sec (the duration of the stimulus). This time period is somewhat arbitrary and it is

unclear over what time period the spike count should be taken. While we made sure our psychophysical studies used the same duration stimulus as the physiological recordings, it has been demonstrated that thresholds for direction discrimination in humans asymptote in around 100-200 msec of stimulus duration (de Bruyn and Orban 1988). It is therefore clearly possible that information is not gathered throughout the whole 1 sec period of stimulus presentation. As duration is lengthened the total number of spikes will increase and the variance will decrease resulting in an increase in the signal to noise ratio (defined by the mean divided by the standard deviation) will decrease allowing better performance. This effect of increasing duration of presentation upon a cell's ability to reliably detect a stimulus configuration has been recently demonstrated in monkey striate cortex (Zohary et al. 1990). Secondly, psychophysical experiments usually employ a technique in which two stimuli are presented and compared; thus the variability in determining the direction of each stimulus must be considered. This contrasts with the normal physiological practice of presenting just one stimulus. We chose to use a psychophysical technique of presenting just one stimulus in order to complement the single cell recordings. However, other researchers (e. g. Bradley et al. 1987) have chosen to compare the responses of a cell to two stimuli closely spaced in time. Under these conditions they found that thresholds (of orientation and spatial frequency discrimination) for individual neurons are improved when compared to similar estimates compiled by the comparison of responses to stimuli which were presented at intervals over a period of several minutes (a situation similar to the present study). This difference is accounted for by the fluctuations in a cell's responsiveness over the course of a few minutes, a well documented finding (e. g. Tolhurst et al. 1981) which adds to

the estimate of a cell's variance. As our data were also collected over a similar time period our calculation of variance, and therefore discrimination, may also underestimate the cells' abilities. On the other hand the reasons for these fluctuations over time are poorly understood, and may have some connection to the preparation of the animal (anesthesia and paralysis) which would be avoided in our experiments.

Peak direction discrimination in these cells (as determined by our analysis) does not occur at the direction to which the cell is tuned (i.e. the one to which it gives the greatest response), but to a direction which is a little away from the preferred direction. While researchers have shown the variation coefficient to be at a minimum at the response peak (Heggelund and Albus 1978) this seems to be outweighed by the steeper rate of change of response evident on the response flanks. A similar result has been reported after analysing the response of cat striate neurons to changes in orientation (Bradley et al. 1987: Scobey and Gabor 1989) and is implicit in the results of Parker and Hawken (1985). If it is assumed that detection of motion of a random dot pattern is mediated by cells which are tuned to the direction in question (e.g. Newsome et al. 1989) then this implies different cells mediate threshold detection and suprathreshold discrimination of movement.

References

Adelson EH, Bergen JR (1985) Spatiotemporal energy models for the perception of motion. J Opt Soc Am A 2: 284-299

Albright TD (1984) Direction and orientation selectivity of neurons in visual area MT of the macaque. J Neurophysiol 52: 1106-1130

Albus K (1980) The detection of movement direction and effects of contrast reversal in the cat's striate cortex. Vision Res 20: 289-293

Baker CL, Braddick OJ (1982) The basis of area and dot number effects in random dot motion perception. Vision Res 22: 1253-1259

Ball K, Sekuler R, Machamer J(1979) Detection and identification of moving targets. Vision Res 23: 229-238

Borst A, Egelhaaf M (1989) Principles of visual motion detection. Trends Neurosci 12: 297-306

Braddick OJ (1974) A short-range process in apparent motion. Vision Res 14: 519-527

Bradley A, Skottun BC, Ohzawa I, Sclar G, Freeman RD (1987) Visual orientation and spatial frequency discrimination: a comparison of single neurons and behavior. J Neurophysiol 57: 755-772

De Bruyn B, Orban GA (1988) Human velocity and direction discrimination measured with random dot patterns. Vision Res 28: 1323-1335

De Valois RL, Yund EW, Hepler NK (1982a) The orientation and direction selectivity of cells in macaque visual cortex. Vision Res 22: 531-544

De Valois RL, Albrecht DG, Thorell LG (1982b) Spatial frequency selectivity of cells in macaque visual cortex. Vision Res 22: 545-559

Dean AF (1981) The variability of discharge of simple cells in the cat striate cortex. Exp Brain Res 44: 437-440

Dow BM, Snyder AZ, Vautin RG, Bauer R (1981) Magnification factor and receptive field size in foveal striate cortex of the monkey. Exp Brain Res 44: 214-228

Downing C, Movshon JA (1989) Spatial and temporal summation in the detection of motion in stochastic random dot displays. Invest Ophthal Vis Sci (Suppl) 30: 72

Gattass R, Gross CG (1981) Visual topography of striate projection zone (MT) in posterior temporal sulcus of the macaque. J Neurophysiol 46: 621-638

Gilbert CD (1977) Laminar differences in receptive field properties of cells in cat primary visual cortex. J Physiol (Lond) 268: 391-421

Gulyás B, Orban GA, Duysens J, Maes H (1987) The suppressive influence of moving textured backgrounds on responses of cat striate neurons to moving bars. J Neurophysiol 57: 1767-1791

Hammond P, MacKay DM (1975) Differential responses of cat visual cortical cells to textured stimuli. Exp Brain Res 22: 427-430

Hammond P, MacKay DM (1977) Differential responsiveness of simple and complex cells in cat striate cortex to visual texture. Exp Brain Res 30: 275-296

Hammond P, Pomfrett CJD (1989) Visual texture: a tool for distinguishing simple from complex neurons in the cat's striate cortex and for elucidating cortical processing. Ophthal. Physiol. Opt. 9: 345

Hawken MJ, Parker AJ, Lund JS (1988) Laminar organization and contrast sensitivity of direction-selective cells in the striate cortex of the Old World monkey. J Neurosci 8: 3541-3548

Heggelund P, Albus K (1978) response variability and orientation discrimination of single cells in striate cortex of cat Exp Brain Res 32: 197-211

Holub RA, Morton-Gibson M (1981) Response of visual cortical neurons of the cat to moving sinusoidal gratings: response-contrast functions and spatiotemporal interactions. J Neurophysiol 46: 1244-1259

Livingstone M, Hubel DH (1981) Effects of sleep and arousal on the processing of visual information in the cat. Nature (Lond) 291: 554-561

Mikami A, Newsome WT, Wurtz RH (1986) Motion selectivity in macaque visual cortex. I. Mechanisms of direction and speed selectivity in extrastriate area MT. J Neurophysiol 55: 1308-1327

Motter BC, Poggio GF (1984) Binocular fixation in the rhesus monkey: spatial and temporal characteristics. Exp Brain Res 54: 304-314

Movshon JA, Thompson ID, Tolhurst DJ (1978) Spatial summation in the receptive fields of simple cells in cat's striate cortex. J Physiol (Lond) 283: 53-77

Nakayama K, Tyler CW (1981) Psychophysical isolation of movement sensitivity by removal of familiar position cues. Vision Res 21: 427-433

Newsome WT, Britten KH, Movshon JA (1989) Neuronal correlates of a perceptual decision. Nature (Lond) 341: 52-54

Newsome WT, Paré EB (1988) A selective impairment of motion perception following lesions of the middle temporal visual area (MT). J Neurosci 8: 2201-2211

Parker A, Hawken M (1985) Capabilities of monkey cortical cells in spatial-resolution tasks. J Opt Soc Am 2: 1101-1114

Regan D, Beverley KI (1983) Spatial frequency discrimination and detection and comparison of post-adaptation threshold. J Opt Soc Am 1684-1690

Robinson DA (1963) A method of measuring eye movement using a scleral search coil in a magnetic field. IEEE Transactions of Biomedical Engineering 10: 137-145

Rose D (1979) An analysis of the variability of unit activity in the cat's visual cortex. Exp Brain Res 37: 595-604

Schiller PH, Finlay BL, Volman SF (1976) Quantitative studies of single-cell properties in monkey striate cortex: I. Spatiotemporal organization of receptive fields. J Neurophysiol 39: 1288-1319

Scobey RP, Gabor AJ (1989) Orientation discrimination sensitivity of single units in cat primary visual cortex. Exp Brain Res 77: 398-406

Skottun BC, Grosof DH, De Valois RL (1988) Responses of simple and complex cells to random dot patterns: a quantitative comparison. J Neurophysiol 59: 1719-1735

Snodderly DM, Kurtz D (1985) Eye position during fixation tasks: Comparison of macaque and human. Vision Res 25: 83-98

Snowden RJ, Erickson RG, Treue S, Andersen RA (1990) Transparent motion stimuli reveal divisive inhibition in area MT of macaque. Invest Ophthal Vis Sci (Suppl) 31: 399

Snowden RJ, Treue S, Erickson RG, Andersen RA (1991) The response of area MT and V1 neurons to transparent motion J. Neurosci 11: 2768-2785

Tolhurst DJ, Movshon JA, Thompson ID (1981) The dependence of response amplitude and variance of cat visual cortical neurones on stimulus contrast. Exp Brain Res 41: 414-419

Tolhurst DJ, Movshon JA, Dean AF (1983) The statistical reliability of signals in single neurons in cat and monkey visual cortex. Vision Res 23: 775-785

van Santen JPH, Sperling G (1985) Elaborated Reichardt detectors. J Opt Soc Am 2: 300-

Vogels R, Spileers W, Orban GA (1989) The response variability of striate cortical neurons in the behaving monkey. Exp Brain Res 77: 432-436

Werner G, Mountcastle VB (1963) The variability of central neural activity in a sensory system, and its implications for the central reflection of sensory events. J Neurophysiol 26: 958-977

Zeki SM (1974) Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. J Physiol (Lond) 236: 549-573

Zohary E, Hillman P, Hochstein S (1990) Time course of perceptual discrimination and single neuron reliability. Biol Cybern 62: 475-486

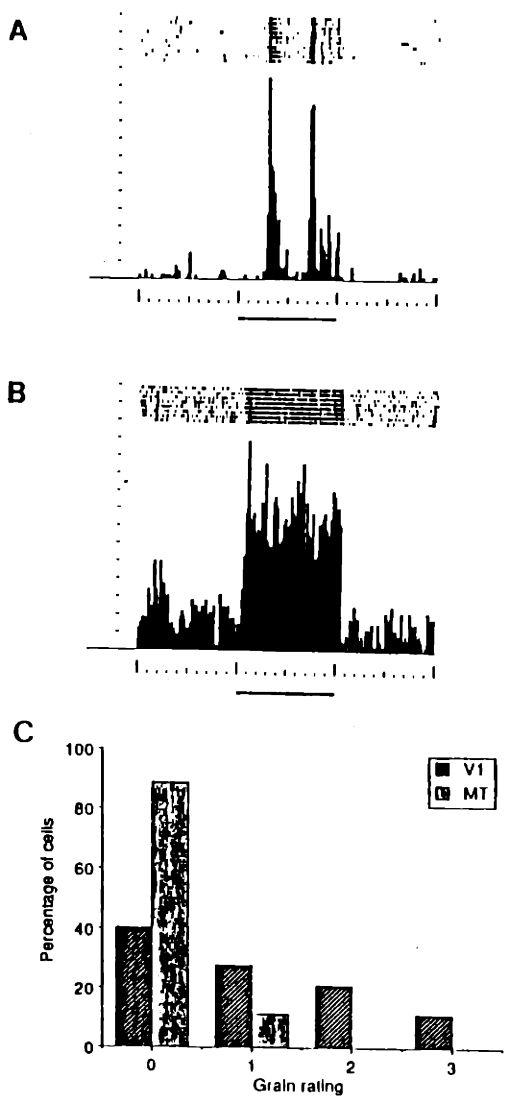


Figure 6.1

Responses of a V1 cell to a moving random dot pattern.
(A) The upper part of the figure shows the response rasters. Each raster indicates a trial, and each dot represents a spike elicited from the cell. Below this is the response histogram constructed by assembling the number of spikes within 20 msec bins and averaging across trials. The presence of the stimulus is indicated by the dark bar under the histogram. The ticks on the Y axis represent 20 spikes/sec/tick, and those of the X-axis 100 msec/tick. The stimulus was of 1000 msec duration. It can be seen that this cell had a tendency to fire at discrete points during stimulus presentation, rather than in a continual manner. It was thus given a grain rating of 3 (see text).
(B) Another V1 cell. This cell fired in an almost continuous manner and was given a grain rating of 0.
(C) Frequency histogram of the type graininess of response in areas V1 and MT.

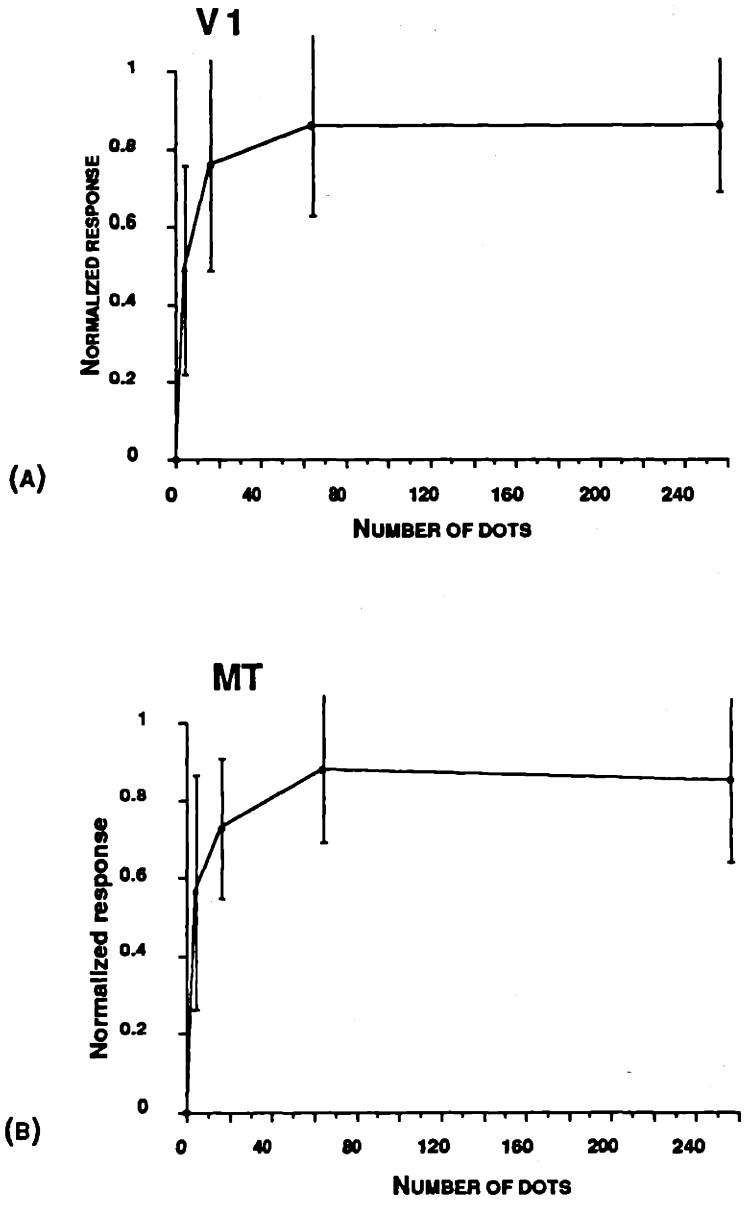


Figure 6.2

Effect of dot density upon magnitude of response in areas V1 (A, N = 37) and MT (B, N = 42). The response from each cell was normalised with respect to its maximum response, and the mean and standard deviation of the population are plotted as a function of the dot density of the pattern.

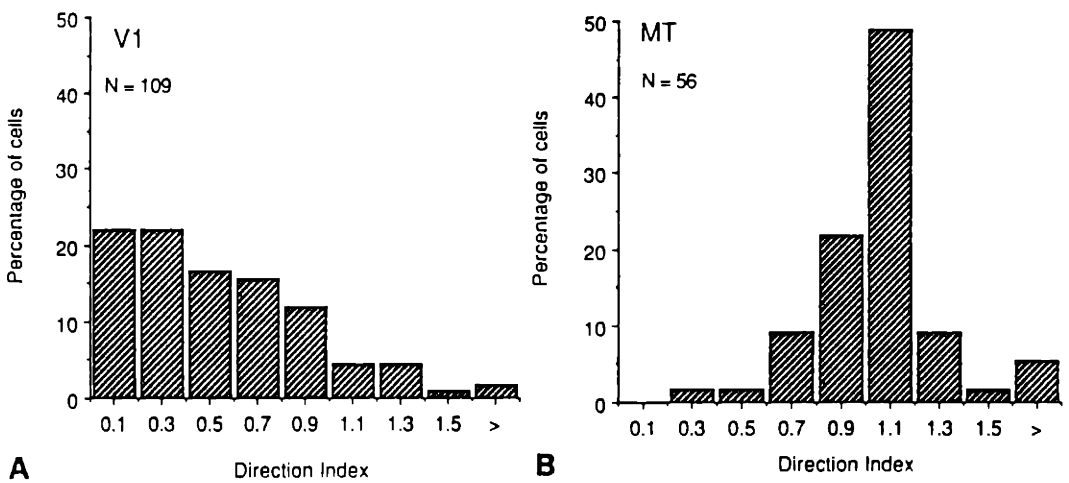


Figure 6.3

Frequency of index of directionality for cells in area V1 and MT.

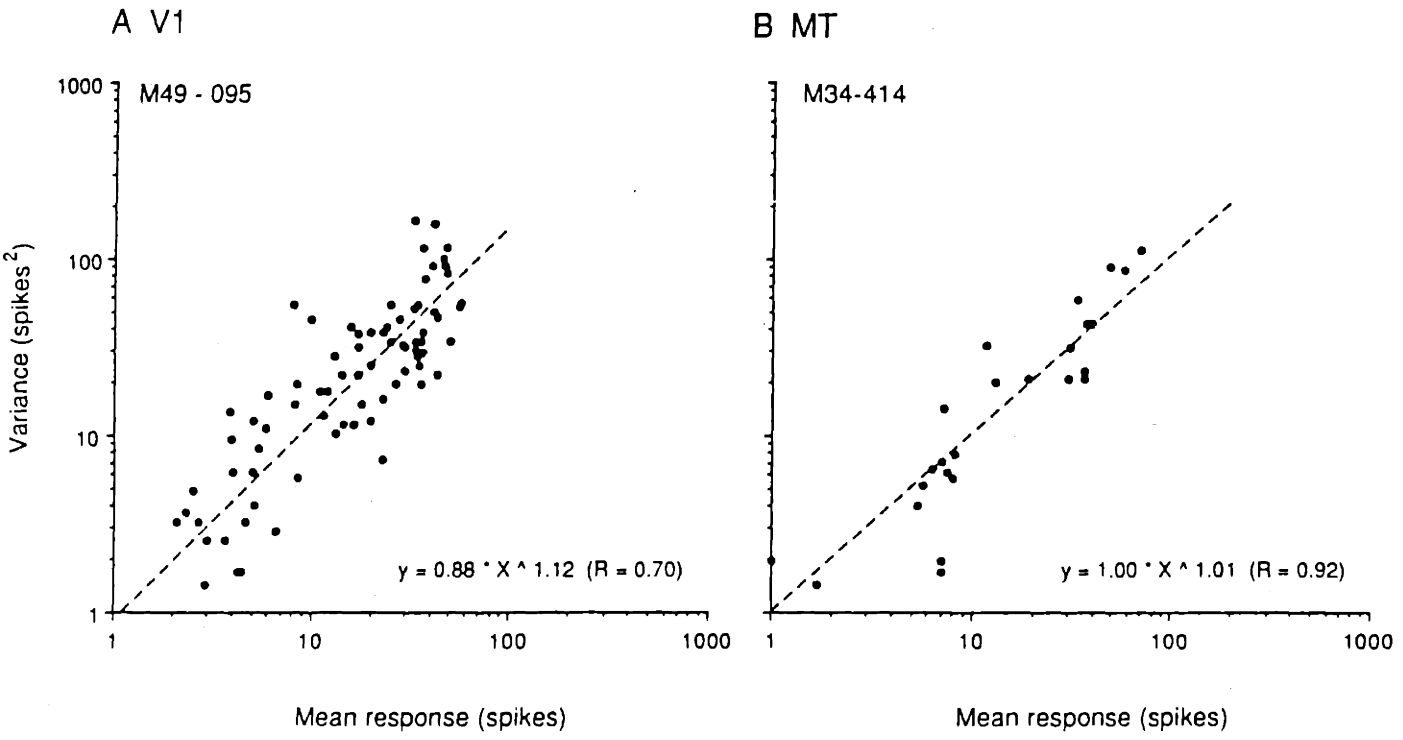


Figure 6.4

Variance as a function of mean response for a cell from area V1 (A) and one from MT (B). The dashed line is the best fitting function of equation 2 (see text), and is indicated at the bottom of the diagram.

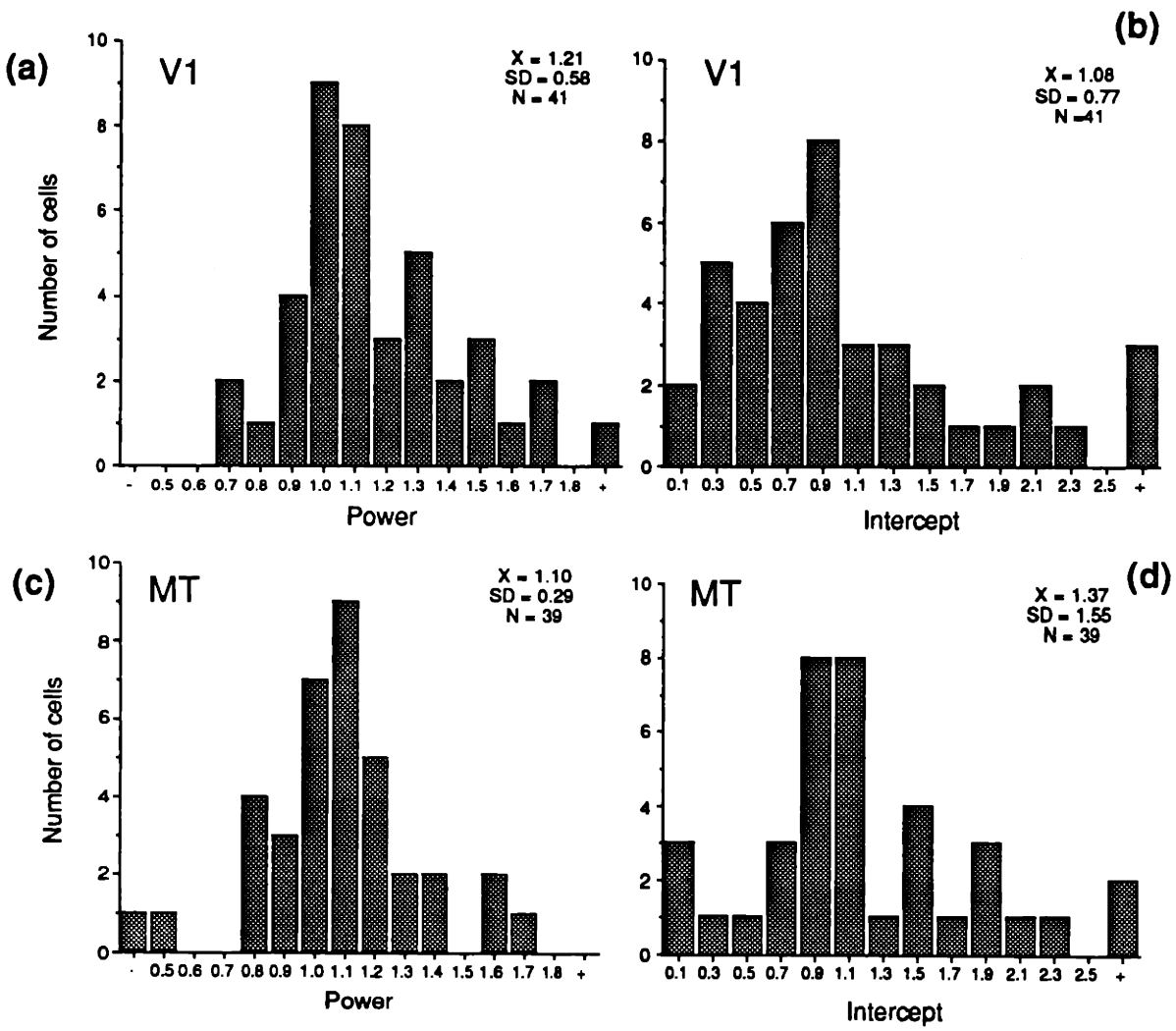


Figure 6.5

(A) Distribution of the slope (power of equation 2) of the fits to the V1 cells.
(B) Distribution of the intercept (constant of equation 2) of the fits to the V1 cells.
(C and D) As for A and B but for MT cells.

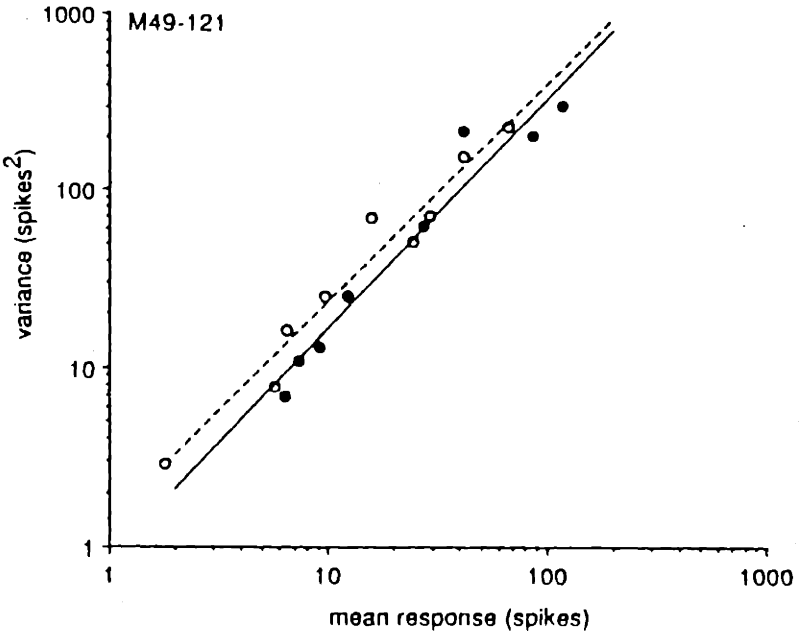


Figure 6.6

As for figure 6.4, but the data points represented by open symbols (dashed line) were produced by varying the speed of the pattern, the solid symbols (and solid line) were produced by varying the direction.

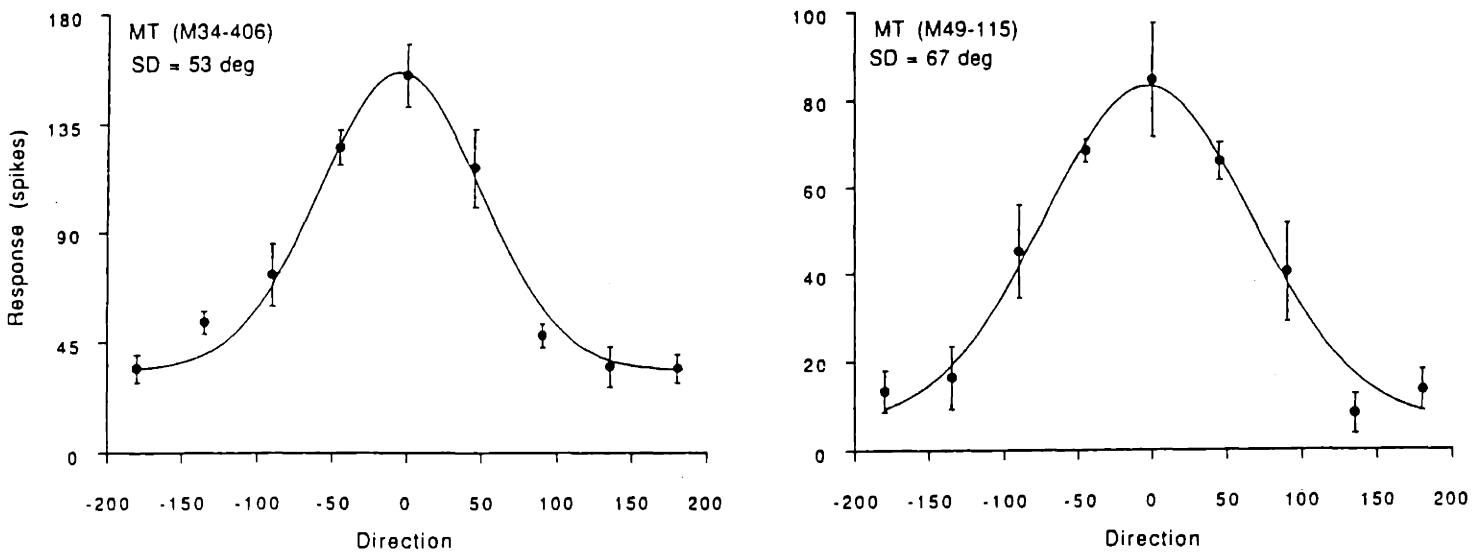


Figure 6.7

Direction tuning for two MT cells. The points represent the mean response, and the error bars the standard deviation. The data were fitted by a Gaussian function (see equation 3), and the standard deviation of the fitted Gaussian is given in the upper left of the figure.

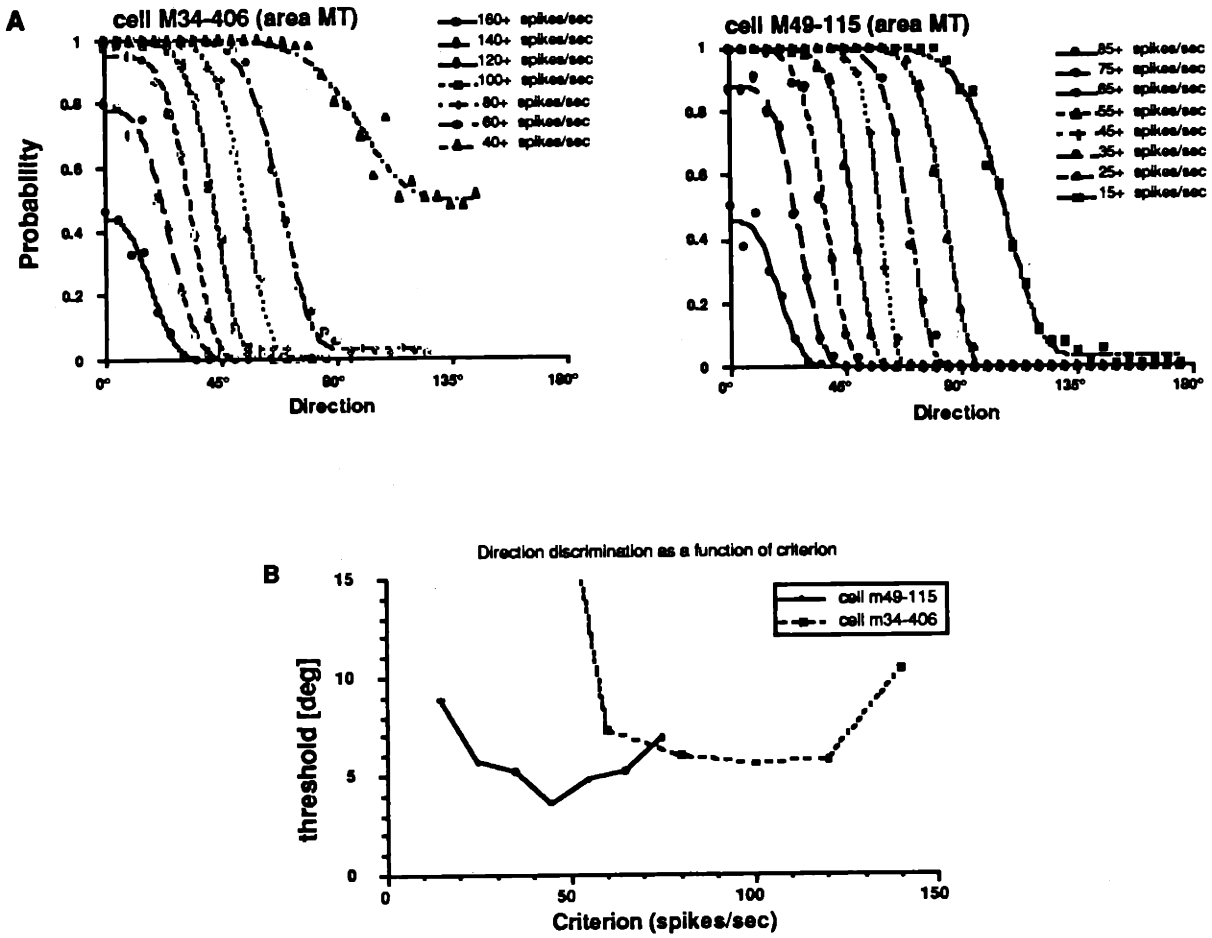


Figure 6.8

(A) Neurometric functions for the two cells illustrated in figure 7. Simulations of the cells to 36 directions of motion were run using our fits to the cells' responses. The neurometric functions were produced by selecting a criterion, e.g. 80 spikes/sec, and calculating the probability that the cell fired at or greater than this rate for each direction of motion. Each plot shows neurometric functions produced by a range of criterion levels. The data were then fitted by an integrated Weibull function.
(B) Direction discrimination for the two cells of A are plotted as a function of stimulus direction. Each data point reflects the threshold derived from one curve in Figure 8a. The x-position for each data point was determined by converting the criterion response into the appropriate direction through the cell's direction tuning curve.

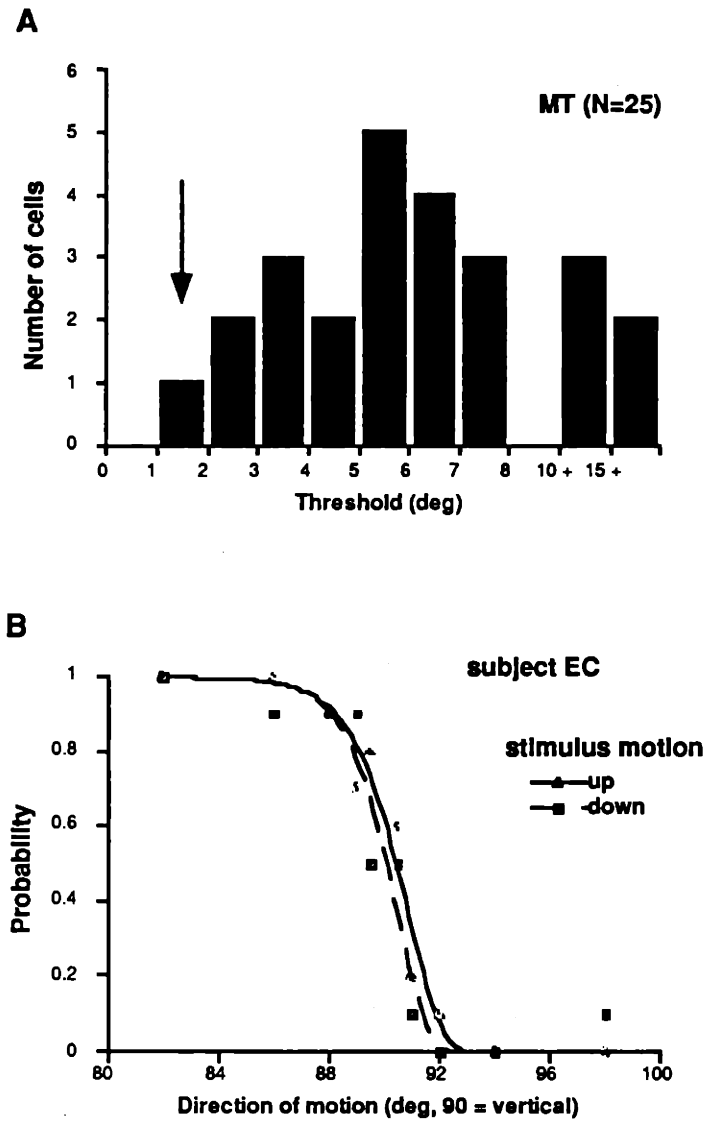


Figure 6.9

(A) Minimum discrimination thresholds for 25 MT cells, determined as illustrated in figure 8. The arrow indicates the mean direction discrimination from a population of 9 human subjects.
(B) Psychometric functions for one human observer. The probability of a 'rightward' response is plotted as a function of direction. The open symbols are for downward motion and solid for upward motion (see Methods). The data were also fitted by an integrated Weibull function.

Chapter 7

MT Beyond Transparency

In this chapter I will discuss some of the issues left unresolved by the work presented in the two preceding chapters. I am currently performing a series of experiments adressing some of the questions posed below and will present some preliminary data.

Do MT cells show double-opponency for direction of motion ?

The two preceding chapters have presented important advances in our understanding of the receptive field structure of cells in primate areas V1 and MT. But in those experiments we had explicitly limited ourselves to what Allman termed the "classical receptive fields" (Allman, Miezin, & McGuiness, 1985a, 1985b). However many cells in the visual system have receptive fields that show an opponency between the receptive field center and the surround. This has been shown for luminance[1], color[2] and motion[3] in a variety of species and visual areas. One interesting receptive field organization is sometimes called double-opponency. This type of receptive field has been found in color-selective cells of the visual cortex. The similarity between the receptive field organization of many luminance, color and motion selective cells suggests that an analogue to the double-color-opponent cell might exist in the motion domain. In fact Frost and Nakayama (1983) claim to have found an even more complex cell type in the optic tectum of the pigeon. These cells are very broadly direction tuned to a stimulus within the classical receptive field but are strongly facilitated when surround is moving in the opposite direction to the center for any direction of motion in the center.

1 as for example in the antagonistic center-surround organization of retinal and LGN cells

2 as for example in the various types of color opponent cells in the LGN and cortex

3 as for example investigated by Allman et al. in the anaesthetized owl monkey

The similarity between the analysis of color and motion in the visual system might even go beyond the existence of double-opponent cells in both and extend to more general similarities in the receptive field structures. Rodieck (1965) for example has demonstrated that the receptive fields of broad-band center-surround cells in the retina can be successfully approximated with DOGs (Difference-of-Gaussians). Most of the color opponent cell types found in the LGN and area V1 can also easily be modeled with such an approach. This might reflect a general principle of receptive field structure in the visual system which could also extend to direction-selective cells. The transformation from color to motion is done by replacing the sensitivity of the cells to opponent pairs of color with sensitivity to opponent directions of motion.

In Figures 7.1 to 7.3 I modeled the responses of 5 real and hypothetical cells to a stimulus moving in the preferred and/or antipreferred direction of the cell for different positions along a line crossing the receptive field (the preferred direction of motion being defined as the direction eliciting the greatest response when stimulating the receptive field center and the antipreferred direction being 180° from this direction).

A center-only cell (Fig. 7.1a) responds to its preferred direction of motion and shows no response to the antipreferred direction. The response of the cell increases with an increasing stimulus size until the cell saturates or the stimulus size reaches the extent of the receptive field.

The center-only directionally opponent cell (Fig. 7.1b) resembles the center-only cell but is inhibited by motion in the antipreferred direction in the receptive field center. This type is very similar to the Type II color-opponent cell which also lacks a surround.

A center-surround spatially opponent cell (Fig. 7.2a) responds similarly to the center-only cell except that the response would actually decrease when parts of the stimulus extend beyond the center of the receptive field. This is because the response of such a cell is the sum of an excitatory input from one simple cell (dotted line) and the inhibitory input from another simple cell with a wider receptive field (dashed line). In its idealized form (as plotted in the figure) the areas under the two curves are of the same size thus making the cell unresponsive to a large uniform stimulus. Note that the two input cells have the same preferred direction.

The center-surround directionally opponent cell (Fig. 7.2b) resembles the center-only directionally opponent cell in that it is inhibited by the antipreferred direction and resembles the center-surround spatially opponent cell in that this inhibition extends into the surround.

Just as the center-only directionally opponent cell can be described as the combination of two center-only cells the double-opponent cell (Fig. 7.3) results from the combination of two center-surround spatially opponent cells. The receptive fields of these two cells are of the same size but their preferred direction are opposite and their input are of opposite sign. Thus the preferred direction of motion of the double-opponent cells is determined by the input cell which makes excitatory connections with the double-opponent cell. A double opponent cell does not only show an opponency between the response of the center and the surround but also shows an opponency within the receptive field center and surround respectively. This opponency means that the addition of antipreferred motion to the center would lower the response of the cell to a stimulus moving through the center in the preferred direction. In

addition the response of a cell to motion of its preferred direction in the center of the receptive field could be enhanced by stimulating the surround with antipreferred motion.

There is considerable circumstantial evidence for the existence of such double-opponent cells in the primate. Allman in his landmark study of the anaesthetized owl monkey demonstrated a surround beyond what he called the classical receptive field (CRF) of MT neurons which could upon proper stimulation lower or raise the response of the cells to stimulation of their CRF. A similar result has recently been reported by Born and Tootell (1992) in the anaesthetized Aotus monkey. Interestingly both studies show that if a MT cell has a surround there always is at least one direction of motion for the surround that will lower the response evoked from the stimulus in the receptive field's center. Evidence for a facilitatory influence from the surround is much weaker. Allman et al. claim that a majority of their cells show such an effect but the median of their population of cells shows no facilitation, even when the surround motion is opposite to the motion in the center. Similarly Born and Tootell claim to see such facilitatory effects but do not show any examples. This weakness or absence might be genuine or it could be an effect of the anaesthetic on the balance of inhibitory and excitatory influences in area MT since both of these studies were carried out in the anaesthetized monkey.

Since the experiments presented here are performed using awake behaving monkeys it will be possible to answer this question. The stimulus used consists of a random dot pattern moving within a round stationary virtual aperture centered on the receptive field and another pattern moving within a stationary virtual annulus also centered on the receptive field

(Fig. 7.4). The direction, density and speed (as well as the sizes of both stimulus parts) can be varied independently. Furthermore 2 random dot patterns moving independently within each of the two stimulus parts can be generated. In the preceding chapters I have demonstrated the power of such transparent stimuli in investigating inhibitory and excitatory interactions within the receptive fields of MT and V1 neurons. In this study the transparent stimuli might prove especially useful since there is evidence that the center and the surround might not always be distinct regions of a cell's receptive field but could reflect different spatial summation for different stimulus directions (Komatsu & Wurtz, 1988).

In a further stimulus variation I am able to mask the annulus used to only stimulate parts of the surround of MT receptive fields (darker regions in the surround in Figure 7.4). It is possible that the area from which the response to a stimulus in the receptive field center can be influenced is not surrounding the CRF but is only on one side of it (Figure 7.5 C). Neither Allman et al. nor Born and Tootell investigated this question but this is an important issue because there are computational approaches to center-surround receptive field organization that assume an asymmetry of the surround (Royden, personal communication). Specifically such a non-symmetric surround would be well suited to detect shear borders (Figure 7.5 D) while a symmetric center-surround organization (Figure 7.5 A) would perform a figure-from-ground detection for small stimuli moving across larger backgrounds.

Many of the earlier studies of center-surround opponency used bars moving across textured patterns. Below I list several disadvantages of this method which we overcome by using transparent random dot patterns.

- Using bars over texture patterns it is very difficult to discriminate between background effects (i.e. interactions within the receptive field center) and surround effects (i.e. interactions between the receptive field center and the surround).
- It is difficult to equate the bar and the texture when trying to quantify inhibitory and excitatory effects.
- Using a bar makes it harder to discriminate between orientation- and direction-selective responses.
- Since the bar is a single element which has to be swept across the receptive field longer stimulus durations as well as the study of sustained responses to continuous motion are difficult. Furthermore speed and sweep duration become confounded.

Preliminary Results

Figure 7.6 shows the response of a typical MT cell with center-surround antagonism. The two left columns of this bar graph show the cell's response to the preferred and antipreferred direction in the center respectively. The third bar demonstrates the strong inhibitory influence that motion in the preferred direction in the center has on the cell's response. The rightmost bar shows that this is one of the MT cells frequently encountered by Allman et al., i.e. it shows no facilitation even when the surround moves in the antipreferred direction of the cell. The few cells with a surround that I have recorded from so far seem to be mostly of this type.

Figure 7.7 on the other hand seems to be a double-opponent cell as defined above. It is not only inhibited by motion in the cell's preferred direction in the surround (compare the first and fifth column) but it is also facilitated by the antipreferred motion in the surround (compare first and last column). Furthermore the cell seems to fulfill all the predictions that can be made from the model of a double-opponent cell presented in Figure 7.3, i.e. the cell's response is reduced when motion in antipreferred direction is added to the center (compare columns 5 and 4) and increased when the same direction of motion is added to the surround (compare column 5 and 6). The opposite is also true, the cells response is increased when motion in the preferred direction is added to the center (compare columns 2 and 4) and decreased when that direction of motion is added to the surround (compare columns 8 and 6).

It should be noted, that the advantage of double-opponent cells over single-opponent cells is not a principal one, but rather an increased sensitivity to stimuli like the ones described in Figure 7.5. For example the response elicited from a single-opponent cell by a small object moving across a background might be small because of non-optimal direction or speed while the response of a double-opponent cell could be much higher. Double-opponent cells are therefore well suited to account for the perceptual 'pop-out' of objects defined by motion. Double-opponent cell also would be much more sensitive to shear boundaries.

Area MT could contain a continuum of neurons ranging from those being inhibited by any direction of motion in the surround to genuinely double-opponent cells. I have not yet collected enough data to determine if cells like the one presented in Figure 7.7 fall into such a continuum or whether recording from

the awake monkey will demonstrate a distinct group of double-opponent cells. Nevertheless the preliminary data presented here seem to suggest the existence of cells in macaque MT that combine the inputs from a variety of cells in a way that renders them maximally sensitive to patterns of motion that differentially stimulate the receptive field's center and surround.

References

Allman, J., Miezin, F. & McGuiness, E. (1985) Direction- and velocity-specific responses from beyond the classical receptive field in the middle temporal visual area (MT). Perception. 14, 105-126.

Allman, J., Miezin, F. & Mc Guiness, E. (1985) Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanisms for local-global comparisons in visual neurons. Annual Review of Neuroscience. 8, 407-430.

Komatsu, H. & Wurtz, R. H. (1988) Relation of cortical areas MT and MST to pursuit eye movements. III. Interaction with full-field visual stimulation. Journal of Neurophysiology. 60, 621-644.

Rodieck, R. W. (1965) Quantitative analysis of cat retinal ganglion cells to visual stimuli. Vision Research. 5, 583-601.

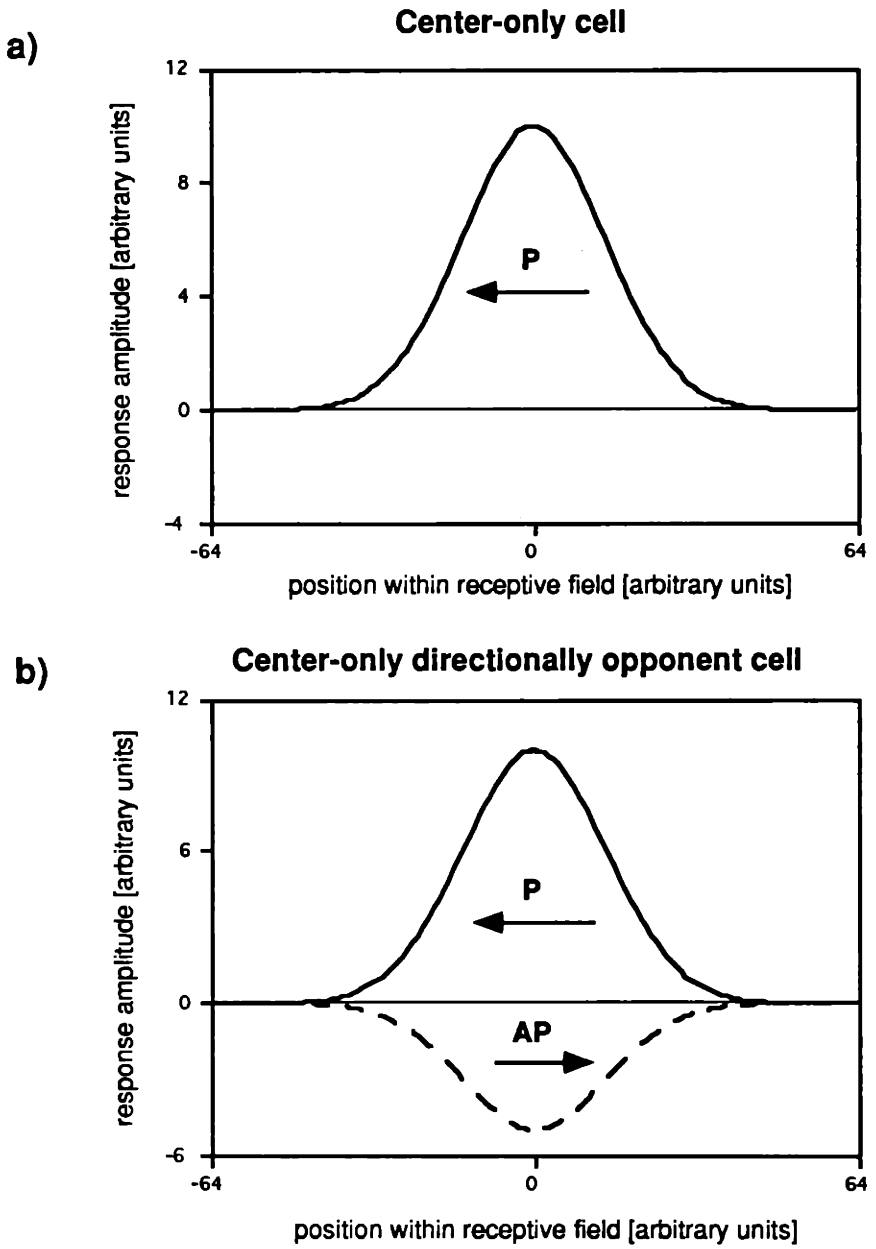


Figure 7.1

Receptive field profiles of a center-only cell (a) and a center-only directionally opponent cell. Curves represent the cell's response to a given direction of motion as a function of stimulus position within the recptive field. See text for details. The preferred direction of motion (P) is arbitrarily assumed to be leftward.

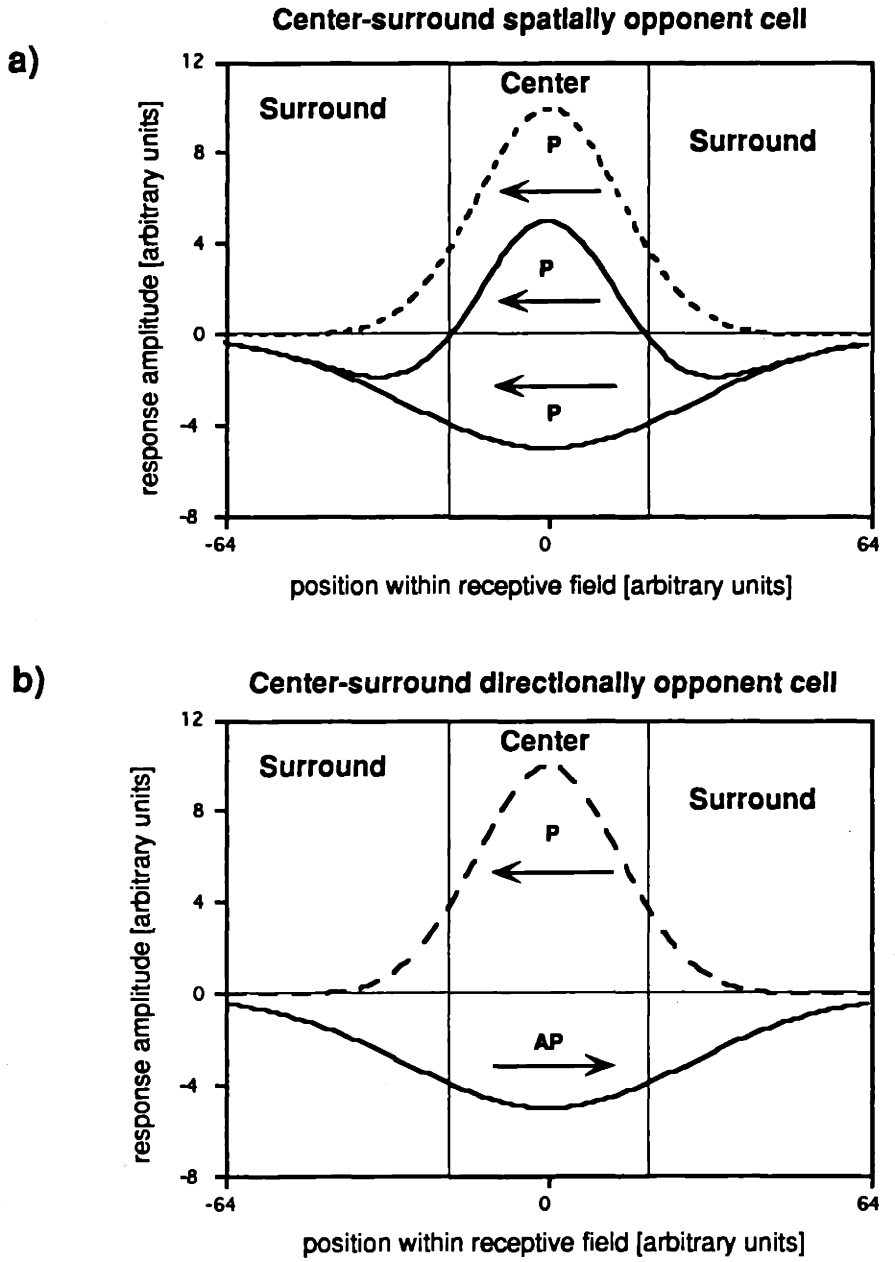


Figure 7.2

Receptive field profiles of a center-surround spatially opponent cell (a) and a center-surround directionally opponent cell. Curves represent the cell's response to a given direction of motion as a function of stimulus position within the recptive field. The middle curve in a) represents the response resulting from the combination of the excitatory and the inhibitory responses. See text for details. The preferred direction of motion (P) is arbitrarily assumed to be leftward.

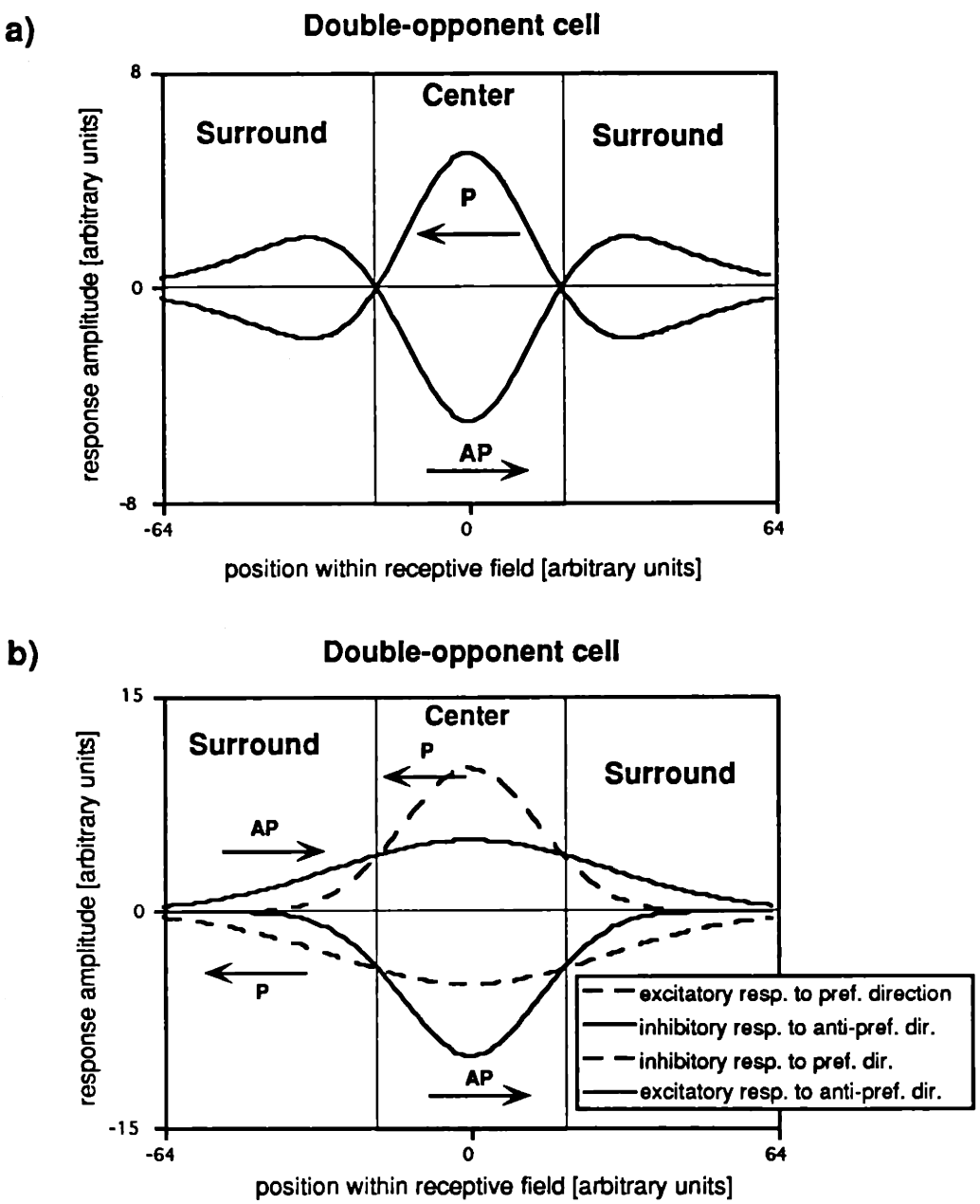


Figure 7.3

Receptive field profiles of a double-opponent cell. In a) the resultant profiles for the two directions of motion are shown while b) depicts the four underlying sensitivitiy profiles. The response for the preferred direction results from the combination of a narrow excitatory response combined with a broad but shallower inhibitory response. The response for the anti-preferred direction on the other hand results from the combination of a narrow inhibitory response combined with a broad but shallower excitatory response. See text for details. The preferred direction of motion (P) is arbitrarily assumed to be leftward.

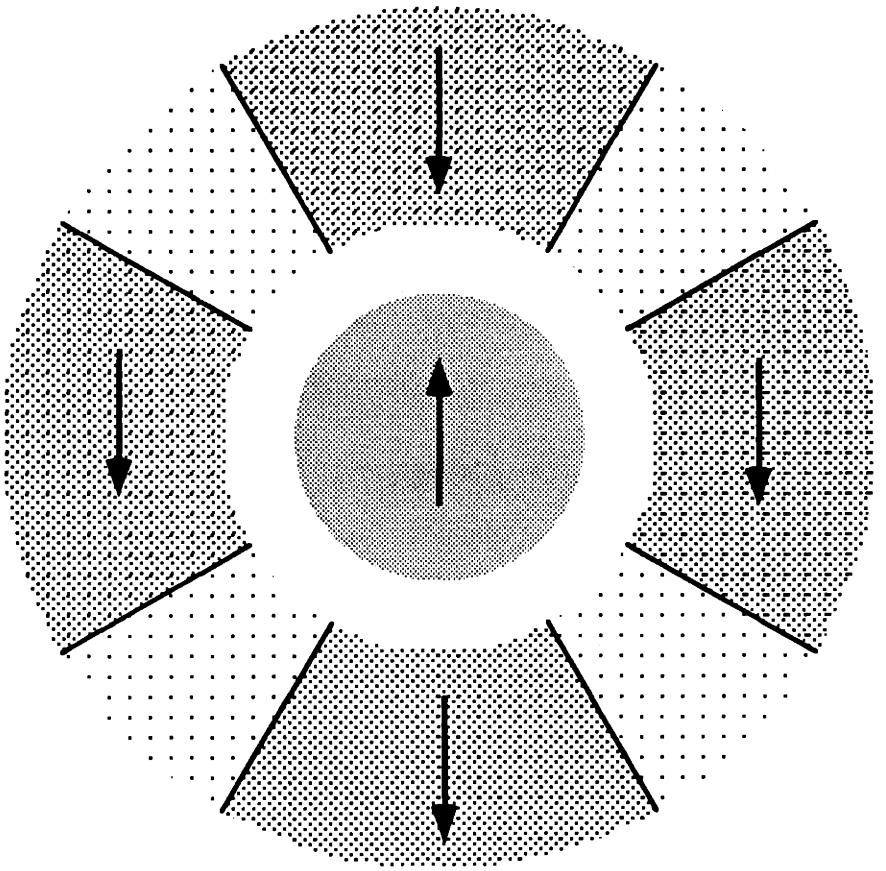


Figure 7.4

Cartoon of the stimulus used. The grey circle in the middle represents the stimulus covering the classical receptive field while the annulus represents the pattern covering the surround. The darker shaded segments of the surround are used in a variation to test for inhomogenities in the surround sensitivity to motion (see text and Figure 7.5). The arrows represent the direction of motion that would be the optimal stimulus for an opponent cell with its preferred direction (as determined in the receptive field's center) being upwards.

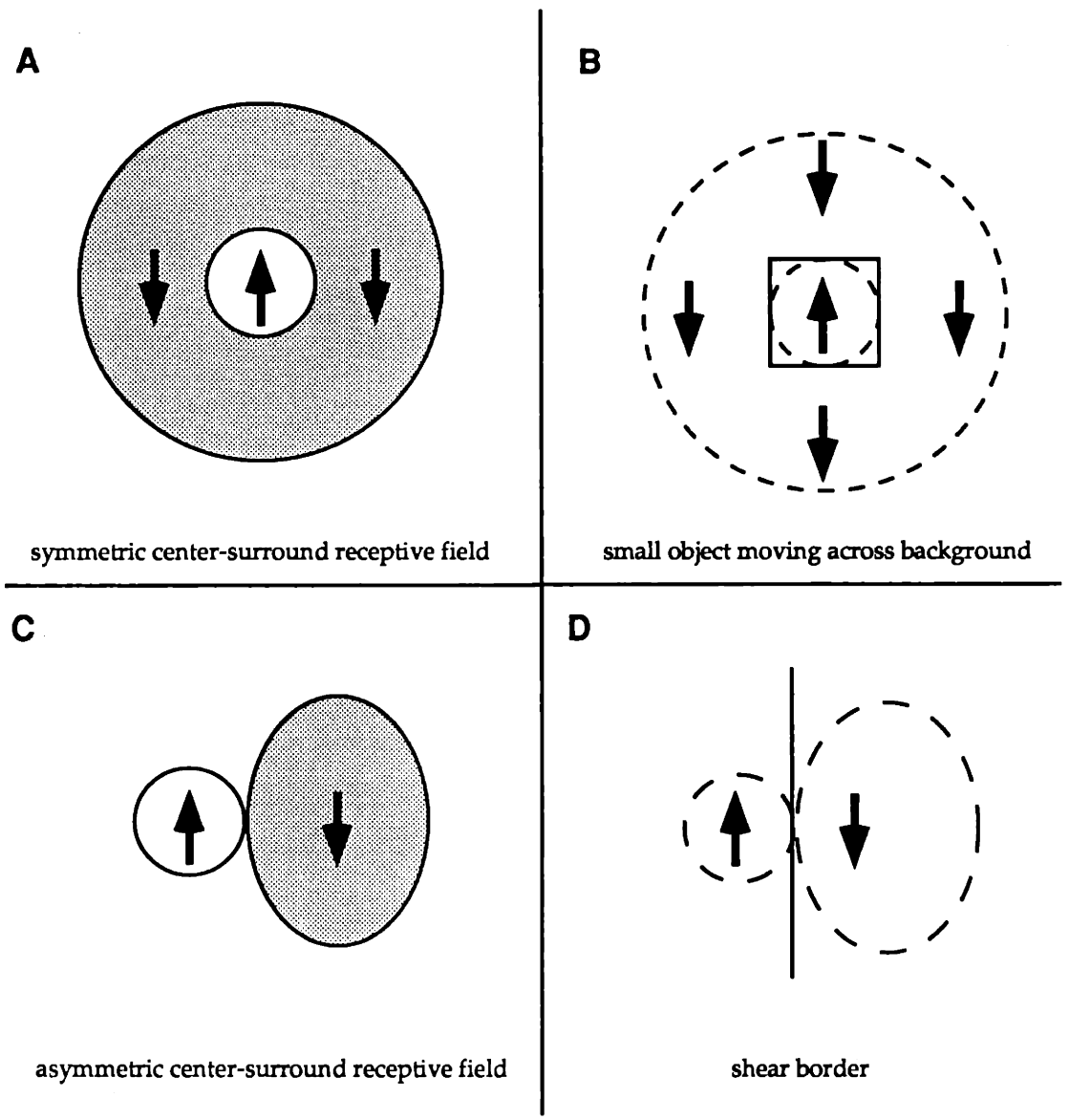


Figure 7.5

Two types of center-srurround receptive field oganizations (**A** and **C**) and their respective optimal stimuli (**B** and **D**). Arrows refer to the preferred directions for the receptive field plots and to the direction of motion for the stimuli. The optimal receptive field position relative to the stimuli is indicated by the dashed lines.

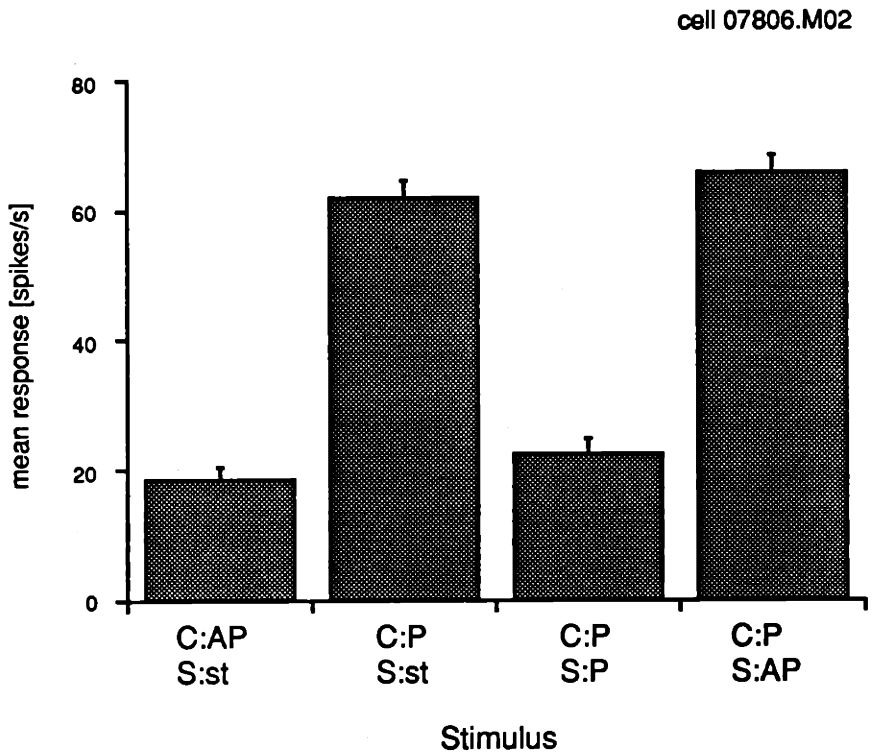


Figure 7.6

Legend for stimulus axis:

C: center stimulus direction

S: surround stimulus direction

P: preferred direction (defined as the direction that elicits the largest response when used in the receptive field center)

AP: anti-preferred direction (defined as the direction opponent to the preferred direction)

st: stationary pattern

Error bars represent standard error of the mean response rate.

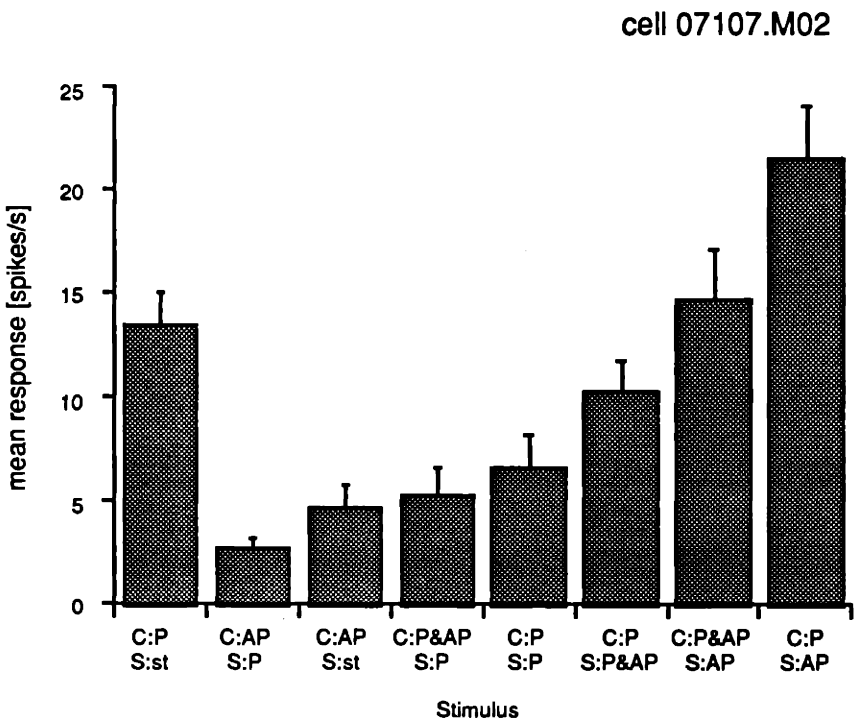


Figure 7.7

Legend for stimulus axis:

C: center stimulus direction

S: surround stimulus direction

P: preferred direction (defined as the direction that elicits the largest response when used in the receptive field center)

AP: anti-preferred direction (defined as the direction opponent to the preferred direction)

P&AP: transparent motion stimulus combining both preferred and anti-preferred direction

st: stationary pattern

Error bars represent standard error of the mean response.