The Analysis of Complex Motion Patterns in Primate Cortex

By

Bard J. Geesaman

A. B. Neurobiology, University of California at Berkeley 1989

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

Doctor of Philosophy

at the

Massachusetts Institute of Technology

September 1995

Signature of Author

Bard J. Geesaman
Department of Brain and Cognitive Sciences
June 29, 1995

Certified by

Richard A. Andersen
Thesis Supervisor
James G. Boswell Professor of Neuroscience

Accepted by

Professor Gerald E. Schneider
Chairman, Department Graduate Committee

To My Friends:

Sandra Lee and Mai-khanh Tran for Caring
Andrew Karson and Hoang Tran for Your Kindness
Paul Loh and Glenn Yamagata for Shared Beliefs
George Daley and Edward Renwick for Your Company

The Analysis of Complex Motion Patterns in Primate Cortex

by

Bard J. Geesaman

Submitted to the Department of Brain and Cognitive Sciences on June 29, 1995 in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Neuroscience

Abstract

I have been interested in how the nervous system processes complex motion patterns such as expansion, rotation, and contraction. The analysis of these types of stimuli is thought to be important for the analysis of self-motion and the motion of objects in the environment. Anatomical, physiological, and psychophysical approaches were all used in an attempt to understand these processes. The four chapters break down as follows:

Chapter 1) This is a review of the literature on cortical area MSTd, a region of the brain thought to be important in the analysis of complex motion patterns. This article starts with an overview of the anatomy and physiology of this area before a discussion of the possible roles this region may play in visual motion processing. The view taken in this discussion is that MSTd is likely to be involved in a number of different functions, possibly including egomotion, object motion, vection, and retinal slip.

Chapter 2) Our lab has proposed that MSTd is important in the analysis of object motion in the environment. Our hypothesis is that this region extracts information about motion pattern from the visual scene and largely ignores the features and cues which define this motion. To test this idea, we recorded from single units in MSTd and obtained tuning curves using motion patterns defined by different features and cues. I found that the preferred complex motion for a particular neuron was independent of the features and cues (i.e. form) of the inducing stimulus.

Chapter 3) Multiple single-unit recording studies in both MT and MSTd have reported clustering of neurons according to preferred tuning. I decided to explore this organization using the double-label 2-deoxyglucose method. The evidence supports a columnar functional organization to both MT and MSTd, with units tuned for expansion and contraction maximally separated in a mosaic of columns of altering specificity.

Chapter 4) There are many more units in MSTd tuned to expansion than to rotation. To see if this anisotropy in response distribution had perceptual consequences, we compared perceived average dot speeds in random-dot complex motion patterns. We found that dots in expansion stimuli appeared to move about 25% faster than the same speed dots in a rotating pattern. The magnitude of this illusion in various variations of this experiment compared well with reported response characteristics of MSTd cells.

Acknowledgments

I thank Richard Andersen and Ning Qian for their support and guidance over the last four years.

I would like to thank the members of the Andersen laboratory for creating such a friendly and stimulating surrounding. I doubt if I will ever work with such a brilliant group of people again, especially if I decide to explore a career in medicine. Special thanks to Catherine Cooper, Steve Marchetti, and Gail Robertson for their technical assistance. Stefan Treue and Ning Qian proved ideal mentors in the area of systems neuroscience. I would also like to thank Pietro Mazzoni for his help and for filling the lab with refined choices of music. David Bradley was invaluable for his patient explanations of statistical issues and overall good taste.

The last four months of my thesis work, I had the privilege to work with Richard Born. If it wasn't for his generous outlay of time and support, Chapter 3 would not have been completed. Thanks Rick for being such a good guy.

I thank Peter Schiller, Ken Nakayama, Ning Qian, and Mriganka Sur for serving on my thesis committee. Also I am indebted to Jan Ellertsen and Jerry Schneider for piloting me through my final year without a local advisor.

I would also like to extend my thanks to my students and colleagues at Quincy House, my home for the past three years. In particular I would like to express my appreciation to George Daley and William Bennett, who have shared their wisdom and friendship. A special thanks to Michael Shinagel, Master of Quincy House, for providing me with a free roof over my head and introducing me to Single Malt Scotch.

Finally, I would like to thank my parents for their continuing support. They instilled in me a love for knowledge and education that has kept me in school far too long.

Table of Contents

Chapter 1

Cortical Area MST and Optical Flow: A Review

Introduction

The purpose of this article is to review what is known about the medial superior temporal (MST) area of the macaque monkey. The first section is dedicated to a discussion of the anatomy and physiology of this cortical area. The remainder of the article is a discussion of the various roles that MST may be playing in visual information processing. Particular emphasis is given to reviewing the literature of ego-motion (direction of heading) representation and pattern motion perception. In addition, we will consider strategies thought to be used by other species to solve these problems. The scope of the discussion will be quite broad, drawing on studies from physiology, psychophysics, and computational neurobiology. In the spirit of systems neuroscience, an attempt will be made to integrate all three approaches. We will propose that the function MST plays is likely to be quite generic and that this region could potentially be involved in many perceptual processes.

Anatomy and Physiology

MST is located medial and anterior to area MT on the floor and anterior bank of the superior temporal sulcus (STS) of the macaque. The long axis of this region is oriented mediolaterally and has a length of approximately 4-5 mm along this direction (Desimone and Ungerleider, 1986; Saito et al, 1986). Anatomical tracer studies have demonstrated connections with visual polysensory areas in posterior prestriate, parietal, temporal, and frontal cortex (Boussaoud et al, 1990). By far its heaviest input is from area MT, anatomically implicating MST in the processing stream believed to analyze visual motion. Heavy lateral connections to area FST in the fundus of the STS have been established, although the significance of these connections and the role that FST is playing in the hierarchy of motion processing has yet to be determined. Other local projections within the STS include reciprocal connections to polysensory association areas (Desimone and

Ungerleider, 1986; Boussaoud et al, 1990). These regions are thought to be involved in the convergence of information from several sensory modalities and facilitate tasks which require integration of these signals. A descending connection from MST to the lateral terminal nucleus of the accessory optical system has been identified (Maioli et al, 1989). The AOS has been solidly established as an important area for the processing of self-motion in birds (Frost et al, 1990) and rabbits (Simpson, et al 1988) and we will later discuss the possible functional significance of this projection in primates. Forward connections from MST include those to posterior parietal cortex (Boussaoud et al, 1990), in particular to areas LIP and VIP. Ongoing investigations by workers in our lab have provided strong evidence that these regions are involved in representing external space in coordinate frames progressively independent from the projection of the optical array onto the retina (Andersen, 1989). Although the perception of motion can be dissociated from that of spatial displacement (for a review see Nakayama, 1985) they are clearly related under normal perceptual conditions and the connection of MST to these posterior parietal regions may well be related to this association. Another forward connection has been established between MST and the frontal eye fields. As will be discussed shortly, MST has been implicated as a step in the sensory-motor loop involved in smooth pursuit eye movements. Given that the frontal eye fields have been implicated in the initiation of eye movements, this projection is conceivably part of this same system. In general, MST gets its inputs from prestriate visual cortex, has intermediate connections with parietal areas, and projects forward to the frontal eye fields and areas in the rostral STS (Boussaoud et al, 1990).

MST was originally thought to be a single functional region based on early anatomical studies, but later investigations with single unit recording suggested that there are at least two or three functional subdivisions (Wurtz et al, 1988; Newsome et al, 1988; Komatsu et al, 1988). Under one established classification, MST has been divided into a dorsal (MSTd), a ventral (MSTl), and an intermediate region (MSTi). Somewhat confusingly, different groups have disagreed about whether to pool data collected in MSTi with that in MSTl or MSTd.

A crude sort of retinotopy has been established across these regions. The most dorsal and ventral sections independently represent the central visual field, while the intermediate region covers the periphery (Tanaka et al, 1993). In terms of functional anatomy, some clustering of response selectivity has been reported by several different groups (Tanaka et al, 1993; Saito et al, 1986; Lagae et al, 1994). Lagae has gone so far as to propose a columnar organization to this selectivity. In his conception, each isotuning column contains units tuned to a single motion pattern, e.g. expansion, contraction, or rotation. According to his findings, response selectivity gradually shifts moving orthogonal to the long axis of the isotuning columns. For example, progressing across these columns selectivity may shift from cells tuned to expansion, to rotation, to contraction, to the opposite direction of rotation, and finally back to expansion. In this scheme, columns for expansion and rotation are maximally far apart, as are the columns for the two directions of rotation. This cycle is thought to repeat every 800-900 microns.

Response Selectivity

Given the widely held belief that both MST and MT are involved in the processing of motion information, it is important to be able to distinguish between the two regions based on physiology (Lagae et al, 1994). One of the most obvious differences between the two areas is that MST has much larger receptive fields than MT. The diameter of an MT receptive field is approximately equal to the distance from the fovea to its receptive field center. There is no such rule for neurons in MST, where receptive field sizes can be as large as 100 degrees across in diameter and whose median size is somewhere between 20 and 40 degrees, depending on the type of stimulus used to map the receptive field. In terms of comparing their responses to motion patterns, both regions respond well to stimuli containing unidirectional (linear, or translational) motion, expansion, contraction, and rotation. However, unlike MST cells, MT cells are not responding to expansion, contraction, and rotation per se but to the local, nearly linear motion signals which make up these patterns. MST cells as a population give significantly weaker responses (in terms of firing rate) to linear motion, expansion, and deformation than MT cells. However, if activity is normalized based on responsiveness to linear motion, MST responds better to rotation and less to flicker, compared to MT (Lagae et al, 1994). The same study reported that, with regards to selectivity along the expansion/contraction axis, MST cells are more directionally selective. This increased selectivity is reflected in the finding that MST cells respond, on average, to fewer motion types than MT cells and that they are more likely to show an inhibitory response to a stimulus in the anti-preferred direction (Lagae et al, 1994). In MT, responses to such elementary flow components (EFCs) as divergence (expansion/contraction), curl (rotation), and deformation depend on the spatial location of these stimuli in the cell's receptive field, but in MST positional invariance to EFCs in a significant percentage of units has been reported by several groups (Lagae et al, 1994; Graziano et al, 1994). Finally, studies probing the subunit structure of units in these two areas have shown that MST cells often have their excitatory and inhibitory regions of their receptive fields overlap, while MT maintains a center/surround spatial separation (Tanaka et al, 1986).

As mentioned above, MST has been divided into different subregions based on differences in response characteristics and proposed function. MSTl and MSTd should be considered as two distinct cortical areas. MSTd prefers large, textured, motion stimuli (Tanaka et al, 1993; Tanaka et al, 1989ab; Saito et al, 1986; Komatsu et al, 1988; Wurtz et al, 1988) and has neurons broadly tuned with respect to speed. Cells respond well to stimuli with speeds as low as 1 degree/second and gradually increase their responses until an amplitude plateau is reached at speeds of about 20 degrees/second (Komatsu et al, 1988). Most of the cells in MSTd are directionally tuned with respect to translational motion, but preferred and anti-preferred directions commonly reverse with stimulus size, indicating that, unlike MT cells, these units are not selective for the spatial-temporal fourier energy of the stimulus (Komatsu et al, 1988).

Probably the most distinguishing characteristic of cells in MSTd is their selectivity for the elementary flow components of rotation, expansion, and contraction (Tanaka et al, 1989ab; Saito et al, 1986; Orban et al, 1992). Single unit recordings in the anaesthetized, paralyzed macaque have demonstrated greater selectivity for isotropic (real) expansion and rotation compared to axial expansion and rotation. With axial expansion, all features in the stimulus move orthogonal to a motion border which bisects the stimulus. The features on the two sides of this border have their motion vectors pointed 180 degrees away from each other. Therefore, in these displays there are only two directions of motion in the image, moving in opposite directions. Axial rotation, which is the same things as shear, has similar motion components as axial expansion, but in this case the velocity vectors on either side of the motion border are oriented parallel to it. This preference for isotropic patterns is interesting, although little attention has been given to it in terms of its possible functional significance.

Early studies in MST that first reported selectivity for EFCs used as a stimulus generator a slide projector with a zoom lens projecting an image consisting of randomly placed dots (Tanaka et al, 1989). The image could be made to expand, contract, or rotate, depending on the proper manipulation of the lens. Several stimulus attributes, in addition to the global motion pattern present in the display, were available for cells in MSTd to potentially respond to. These included a radially oriented speed gradient — the speed of each feature (dot) in the stimulus was a linear function of its distance from the center of the stimulus. Additionally, when zooming the lens to create the expansion pattern, each dot changed its size in the display, another possible cue that MST cells could be selective for, independent of global motion. Finally, the features in all these stimuli had an acceleration component to their motion vectors oriented along the radial axis of the image. It was shown that none of these cues in isolation, with the exception of global motion, gave a tuned response in the cells recorded from (Tanaka et al, 1989). Speed gradient, in conjunction with the spatial arrangement of velocity vectors in the image, had a facilitory role in response selectivity, although the effect was modest.

Positional Invariance

To say that a unit is selective for a global motion pattern such as expansion or rotation, it is necessary to demonstrate that this property emerges independently from selectivity to translational motion. For example, by appropriately placing a subregion of an expanding stimulus over the receptive field of a unit tuned specifically to translation, a brisk response can be elicited that exhibits all the properties of being directionally selective. However, moving the stimulus within such a neuron's receptive field will change the selectivity of the unit, even reversing it, so that it now prefers the opposite type of global motion. So, to say that a unit is selective for a particular EFC, we must demonstrate that its response is positionally invariant with regards to the spatial placement of the EFC within the receptive field. The invariance established at the two spatial locations need not be reflected in the width or the height of the tuning curve, only with regards to the direction of maximal response, i.e. the preferred motion pattern. Approximately half of the cells in MSTd demonstrate this type of selectivity (Saito et al, 1986; Tanaka et al, 1989; Lagae et al, 1994) and this invariance is a distinguishing characteristic when identifying this region based on response properties. Tuning invariance has also been observed for MSTd cells with respect to speed (Lagae et al, 1994) and with respect to cue, feature, and form (see Chapter Two of this thesis).

Single, Double, and Triple Component Selectivity

Several investigators have characterized cells in MSTd based on the number of global motion types (EFCs) the cell is selective for. Selectivity, used this way, is defined as a unit responding significantly differently to stimuli on opposite sides of the stimulus space. For example, we say that a cell is selective for translational motion leftward if the cell fires significantly more when presented with leftward motion than rightward motion. Selectivity in MST can be determined separately for linear motion, divergence (expansion/contraction), and curl (rotation). In this way, units are characterized as being either single, double, or triple component, depending on the number of motion types to which these units are sensitive. To a rough approximation, the population of cells in MSTd is equally distributed into thirds when divided in this way (Duffy and Wurtz, 1991b). Single component cells respond exclusively to either expansion, contraction, one direction of rotation, or translation (linear) motion. Cells of this class exhibit the greatest direction selectivity, are most likely to show significant inhibitory responses, have the greatest degree of positional invariance with respect to stimulus placement in their receptive fields, and have the greatest tendency for their excitatory and inhibitory subregions to overlap (Duffy and Wurtz, 1991b).

There is a significant amount of disagreement about the response characteristics of double component cells. Graziano, et al (1994) found cells selective to the full range of EFC combinations, while another group found no evidence of cells selective for translational motion in combination with any other EFC (Duffy and Wurtz, 1991a). Lagae, et al (1994), agreed with Graziano, et al (1994) with regards to the existence of cells selective both to translational motion and an EFC, but found no cells selective for either expansion/contraction paired with a rotation. This disagreement is an important one because our lab has argued that the presence of this type of double component cell really represents a population of cells tuned to intermediate directions of global motion that appear perceptually as spirals (Graziano et al, 1994). As will be discussed below, the existence of cells tuned to spiral motion patterns is relevant to the way MST may be processing optical flow information. At the present, we have no way of reconciling these conflicting claims.

The triple component units, those cells responding to three types of global motion, as a population had the least amount of directional selectivity, the smallest likelihood of having significant inhibitory responses to anti-preferred stimuli, the least positional invariance, and exhibited the least likelihood of having their excitatory and inhibitory subregions overlap (Duffy and Wurtz, 1991ab). Statistical analysis of response trends between single, double, and triple component cells indicate that this distinction is an arbitrary one. A continuum of selectivity for various combinations of motion types exists in the population of MSTd cells, and no clustering occurs in the parameter space defined by component responsiveness.

Graziano, et al (1994) developed this idea in depth and proposed that rather than multiple-component cells being independently selective for multiple component types, there is one "preferred" global motion stimulus for each cell. For example, consider an observer moving forward through the environment while making a smooth pursuit eye movement leftward. The global motion pattern projected onto his retina will be determined by the vector addition of a rightward linear motion field due to leftward eye rotation and an expanding motion field produced by forward translation. A subpopulation of multiple-component cells will be maximally selective for this particular complex vector field. However, this selectivity is not complete, and presentation of the component motion fields alone (translation or expansion) would also activate this cell, although not drive it maximally. Although selective for a single complex motion pattern, such a unit would be classified as double component.

An EFC that we have not discussed much is deformation. Very few cells in MST have been found that are selective to this type of motion (Lagae et al, 1994). As will be discussed later, the absence of such units argues against this region playing a direct role in analyzing structure-from-motion.

Figure vs. Field Cells

Some reports have divided MST cells into those with "figure" selectivity and those with "field" selectivity (Saito et al, 1986; Tanaka et al, 1989). These studies contain the first discussion of the optical flow vs. object motion distinction that we will explore in greater depth later in this review. This distinction was made using translational motion only, not exploring figure vs. field selectivity with stimuli containing other EFCs such as expansion and rotation. Figure cells were characterized as preferring the translational movement of small dots or edges across the cell's receptive field. Movement of extended fields of dots in such cells elicited poor responses. Field cells exhibited opposite characteristics and frequently reversed their preferred tuning direction for translational motion with changes in stimulus size. It is not clear whether this figure/field distinction reflects two distinct populations of cells, or whether, like the component number classification previously discussed, is arbitrary and there is a continuum of selectivity. There is evidence that the cells in these two classes are not distributed evenly throughout MST and that field and figure cells are primarily localized in MSTd and MSTl, respectively (Komatsu et al, 1988).

Disparity Tuning

A property of MSTd cells that has received recent attention is their disparity tuning (Roy and Wurtz, 1990, 1992). 90 percent of cells tested in MSTd were sensitive to the disparity of the moving stimulus. Like cells found in MT, this disparity tuning was generally relative to the plane of fixation, indicating that these cells are not sensitive to absolute depth. Surprisingly, 95 percent of these cells were near/far cells, as described by Poggio, et al (1988), being broadly tuned for either far or near disparities. This is in contrast with studies in MT which showed a much higher proportion of cells with "tuned near" and "tuned far" selectivity. Apparently, there is a convergence of disparity channels in the projection from MT to

MST, accounting for this partial loss of selectivity. However, some specificity is clearly maintained, unless the broad disparity tuning found in MST is recreated de nova in this region. Because of the broad disparity tuning of these units, this area cannot be involved in the fine disparity system, used in such behaviors as threading a needle. The observed tuning characteristics are consistent with those postulated for the "coarse" disparity system involved in such activities as guiding vergence eye movements. 60 percent of the MSTd cells tuned for disparity give differential responses (in terms of amplitude) to linear motion depending on the disparity of the invoking stimulus, but maintain their selectivity for a single motion direction across the different disparities. However, 40 percent of the cells had their directional selectivities reverse, a finding thought to be relevant (see below) with regards to the possibility that this area is registering self motion through the environment.

Smooth Pursuit

MST has been implicated as one of the nodes of the smooth pursuit control loop. Many cells preferentially fire during pursuit in a particular direction, even across a dark background. Because the great majority of cells in this region don't start firing until after pursuit behavior has begun, it is unlikely that this area is involved in the initiation of pursuit activity (Newsome et al, 1988). However, both the initiation and maintenance of pursuit have been shown to be impaired following lesions to the STS known to include MT and MST. In the studies reporting this deficit, the monkey is trained to initiate pursuit with a saccade to a moving target. It is a deficit in representing this target motion that is thought to account for the impairment of pursuit initiation and not a direct motor impairment. Consistent with this, saccades to stationary targets were unaffected. Therefore, there was no deficiency in representing information about relative position, as the deficit was specific to making saccades to moving targets. As discussed above, this information about target location in space is likely represented in posterior parietal cortical regions, where lesions are known to impair saccades towards stationary targets. With regards to pursuit maintenance, the deficit following lesions to MT/MST was an inability to adequately maintain accurate target fixation. The systematic lag of pursuit behind target motion suggests that the gain of the system registering retinal slip was compromised by these lesions. In these cases, catch-up saccades in the correct direction were made, offering further evidence that the system encoding retinal positional error remained intact (Yamasaki and Wurtz, 1991).

The functional distinction between MSTl and MSTd is particularly evident when considering their respective roles in smooth pursuit. Microstimulation of MSTl, but not MSTd, produces an acceleration of pursuit towards the side of the stimulation (Komatsu and Wurtz, 1988). Units in the two subregions also respond quite differently during periods of pursuit when the target is temporarily extinguished. During smooth pursuit, if the target disappears for short periods of time (200 msec), pursuit can be successfully maintained through the gaps. Interestingly, MSTd cell with pursuit activity continue to respond during this "wink" but MSTl units decrease their activity during the same interval (Newsome et al, 1988). Although this could be explained on the basis of greater temporal averaging in the responses of MSTd cells, experiments where the target was stabilized on the retina provided similar results. Based on these findings, it has been proposed that MSTl has an on-line role in the smooth pursuit control loop: that of detecting the retinal slip of the target and sending this information forward to a motor center so that compensation can be made for target motion. MSTl may partially share this task with area MT, as both areas display similar response characteristics during smooth pursuit. It is believed that this error signal is registering target motion per se, rather than positional error, which is carried out by a separate saccade system in parietal cortex. Possibly, projections from the STS to the dorsal lateral pons (DLPN), an area implicated in the motor aspects of pursuit, may be important for transmitting this pursuit signal.

MSTd has been postulated as being somewhat outside of this loop, with its function limited to registering the consequences of pursuit activity. Given that these cells continue to fire during periods when no visual stimulus is providing an input (during "winks"), the smooth pursuit signal must be extra-retinal, perhaps an efference copy of the motor signal used to drive the pursuit. The cells in MSTd with this property also possess a directionally tuned visual response for large field motion with a selectivity opposite to that of the preferred pursuit direction. These two signals therefore summate when pursuit of a target occurs over an illuminated, textured background, demonstrating ideal response characteristics for a cell presumed to be involved in monitoring pursuit (Wurtz et al, 1988; Newsome et al, 1988; Komatsu et al, 1988). The authors of these studies proposed that this pursuit related activity is sent on to higher cortical centers like 7a for further processing.

Possible Functional Roles of MST

Now that we have covered the fundamental anatomical connectivity and physiological response profiles of MST neurons, we will discuss the possible functional roles that these cells might play. Our discussion will mostly be limited to a consideration of MSTd, as this area has been the most studied. As a starting point, we will take for granted that these units' preference for large translational and radial stimuli is a distinguishing feature that needs to be taken into account when considering the function of this region. The extra-retinal smooth pursuit signal will also need to be taken into account before a complete functional description is made. An inherent conceptual difficulty of this approach is that any particular behavioral or perceptual phenomenon under consideration may not be under the control of any one region of the brain but may be the consequence of several regions working together. Also, it is very likely that MST is involved in multiple perceptions and behaviors that we think of as being functionally distinct. With these caveats behind, let's move forward.

Optical Flow

Since its discovery about a decade ago, MSTd has been thought to play an important role in processing optical flow. This proposal was motivated by the large receptive fields of these neurons as well as the "flow-like" patterns that these units are selective for. This selectivity has generally been considered in the context of using these large-field motion patterns to obtain information about observer movement in the environment, such as direction of heading (DOH). We will see that the processing of optical flow information is relevant for other neural phenomena as well, such as vection, postural stability, and structure from motion. Because of this, we will first provide a brief introduction to optical flow independent of the uses primates make of these stimuli.

Optical flow can be defined very generally as spatial-temporal changes in the optical array and how these changes evolve over time. In this formulation, the source of these changes, whether observer movement or movement within the environment, has intentionally been left out. The optical array is not, as is often believed, the projection of the three dimensional world onto the retina (this is properly referred to as the retinal array). The optical array is a formal convention that is independent of observer orientation. Similarly, optical flow is independent of observer motion other than translation through the environment. Importantly, the optical array and optical flow are not effected by either head or eye rotations. For convenience, we can think of the optical array as the two dimensional projection of the visual world onto any reference surface that is independent of eye and head rotation, but depends on the translational movement of the observer through the environment. Traditionally, the surface of a sphere, whose center is the nodal point of the eye, has been used as this reference. The retinal array and retinal flow are influenced by eye and head rotation, because the reference plane in these two cases is the retina which, being fixed relative to the orbit, moves with the eye and head. Information about the three dimensional structure of the environment, as well as observer translation relative to this environment, is available from the optical flow (at least theoretically), as we will see shortly. In addition, important information about observer eye and head movements is present in the retinal flow. When the eye and head are stationary, optical and retinal flow patterns are identical. With rotation, a uniform translational motion field is added on top of the optical flow pattern to produce the retinal flow. The problem is that only the latter signal is directly available for the nervous system to process, and it is not immediately obvious how to separate out the different components of motion present, so that this information can be used in a behaviorally relevant way. This task has been formalized as involving a decomposition of retinal flow into an exterospecific (containing information about the structure of the environment) and propriospecific (containing information about observer eye and head motion) components (Koenerink and van Doorn, 1981).

The behavioral effects of optical flow have been well established, including eliciting the optomotor response, compensatory postural adjustments, head bobbing in pigeons, and locomotion in lobsters (Pailhous et al, 1990; van Esten et al, 1988; Fluckiger and Baumberger, 1988; Gielen and van Asten, 1990). There is much evidence that the geniculo-striate pathway is important for transferring this information along to extra-striate visual areas like MT and MST, although evidence from cortically blind patients offers support for an alternative colliculus-pulvinar-parietal pathway (Mestre, 1992). Another question is the form in which the nervous system represents flow information. In the computational and psychophysical literatures, flow has been represented as a positional velocity field, with the motion of each image feature encoded with the proper instantaneous direction, position, and speed. Under natural conditions, with time this velocity field generally changes, depending on alterations in the environment and in observer motion. We say that when the velocity field changes over time that the image is evolving and consequently components of acceleration as well as other cues are added to the visual stimulus. For computational, psychophysical, and physiological purposes we often want to be able to produce stimuli where the instantaneous velocity field is presented without additional cues, in order to study its processing in isolation. This can be done, for example, by presenting the system under study with just two successive image frames, as three frames are obviously required to represent information about acceleration. Alternatively, with the random dot displays used in our lab to study MSTd (Graziano et al, 1994; see Chapter Two of this thesis), we simulated the approach to a vertical plane but removed all cues except the positional velocity field. Components of acceleration were eliminated by displacing the dots every frame in accordance with the same vector field. Although the velocity field produced is consistent with an approaching plane, because the velocity field does not change with time the plane remains a fixed simulated distance away from the observer.

If the nervous system only has access to the velocity field and does not process higher order information such as acceleration, or if such a system is deprived of these cues in the stimulus, ambiguities in the optical flow arise. The general belief that the human visual system is poor in detecting acceleration makes the former situation highly relevant. One such ambiguity that has received considerable attention is that the instantaneous velocity field produced by rectilinear (straight) translation in the presence of eye rotation is identical to curvilinear motion with the eyes stationary. Being able to resolve this ambiguity is important behaviorally; therefore information about acceleration is needed to make this distinction based on flow information alone (Warren et al, 1991; Longuet-Higgins, 1986). Alternatively, other modalities (vestibular input, an extra-retinal signal for eye movement) could also resolve these ambiguities.

Patterns of retinal flow can be quite complex, particularly in animals not limited to a terrestrial setting. These types of patterns are all effective stimuli for units in MSTd. The simplest case occurs when the observer makes a head or eye movement with no observer translation and no motion in the environment. In this situation, the retinal flow is a homogeneous array of velocity vectors all with the same orientation and length. The pattern of flow is completely independent of environmental structure, including relative depth. Recti-linear translation parallel to the line of sight, in the absence of eye rotation, produces either expansion for forward motion (all velocity vectors pointed away from the direction of heading), or contraction for backward motion. The spatial arrangement of the velocity vectors, as well as their orientation, is independent of environmental structure and is determined entirely by observer translation. However, the lengths of these vectors varies depending on the layout of the environment, e.g. those vectors closer to the observer have greater magnitude than those further away, for the same retinal eccentricity. For the special case of recti-linear movement orthogonal to the line of site (with no eye rotation), a translational pattern of flow similar to that described for eye rotation results, but the magnitude of the vectors is a function of environmental layout.

Under natural conditions, we generally track features in the environment while moving forward. If the feature is exactly in the direction that we are heading, or if the feature is very far away (the horizon), no eye rotation is required to maintain fixation and a pure expansion flow pattern is produced. However, if the feature we are fixating is nearby (a mark on the sidewalk in front of us), we induce a retinal flow pattern determined by adding the motion vectors that would have been produced separately from the translation and the rotation. Adding the radial flow field to the rotational flow field produces a spiral pattern of flow, assuming some depth variation is present in the image. The exact pattern of the spiral (ratio of expansion/contraction to rotation) is determined by the structure of the environment. As the depth variation in the environment approaches zero, the instantaneous velocity field approaches that of a pure radial (expansion) pattern, appearing like a pure observer translation. Because of this, we have a situation where the optical flow can be ambiguous if the image flow is not allowed to evolve with time or if only the instantaneous velocity field is considered. Finally, because our movement trajectories are commonly not perfectly straight, we often produce hyperbolic flow patterns during periods of curvilinear motion. Even more complex flow patterns are possible, as when, for example, a monkey is swinging through a tree, moving his eyes and head, all independently. Although it is mathematically trivial to predict these flow patterns by adding the different motion contributions separately, going in the opposite direction is much more difficult, and even ill-posed, especially when higher order information such as acceleration is deprived. When there is movement in the environment, such as a tree blowing in the wind, this further distorts the flow pattern, at least locally. We can see that there is a tremendous amount of potential information present in the global motion pattern falling on the retina, but it is a nontrivial task to disentangle all this conflicting information.

Direction Of Heading Determination

Gibson (1950) was the first to recognize that information about the direction of observer translation could be determined from an examination of the optical flow. He correctly identified the global focus of expansion in these patterns as corresponding to the direction of heading. This is true for all recti-linear movement, independent of whether or not there are simultaneous rotations of the eye or head. Unfortunately, the nervous system does not have direct access to optical flow, and the focus of expansion with respect to retinal flow, to which the observer does have direct access, only corresponds to the direction of heading when there is no simultaneous eye or head rotation. Because MSTd responds to the types of flow patterns produced by observer translation, this region has been implicated as the location where direction of heading computations are processed (Judge, 1990). Many workers in MST have labeled it the "ego-motion" center, and we will address the strengths and weaknesses of this proposal before considering alternative roles that this cortical area might play. For convenience, we will use the label ego-motion to refer specifically to direction of heading determination and distinguish this function from related issues such as the sensation of vection and the motor compensations related to ego-motion such as postural stabilization. These additional phenomena will be considered subsequently.

The determination of direction of heading from optical flow has received considerable attention over the past 15 years, particularly in the computational and psychophysical literatures (Prazdny et al, 1980; Poggio et al 1991). Far less attention has been given to the issue of ego-speed, possibly because this parameter is unavailable from optical flow in the absence of information about the absolute depth of environmental features (Larish and Flach, 1990). The problem of determining direction of heading can be formalized as follows: The observer has six potential degrees of freedom for movement, three rotational and three translational. Alternatively, Chasles's theorem states that every movement can be decomposed into a translation and a rotation around a single axis in a unique manner. These two formulations are formally equivalent; four degrees of freedom are used to determine the proper axis in the latter description of the problem. Cues from multiple modalities, including vestibular, auditory, and somatosensory, may be instrumental in facilitating the determination of these parameters, but this discussion will be limited to information provided by retinal flow, perhaps in combination with an extra-retinal signal about eye rotation. Furthermore, in the analysis that follows, we will assume that any rotational component (about the vertical axis) of the flow is a consequence of eye movement and that the head is fixed relative to the shoulders. All the mechanisms discussed that consider an extra-retinal signal about eye rotation could trivially be extended to include an efference copy signal for head rotation.

As alluded to previously, the task of determining direction of heading for the case of recti-linear translation without eye or head rotation is trivial, at least conceptually: identify the global outflow of expansion. Whether or not the nervous system has components capable of detecting this feature in the flow is another question—and this may not be necessary as other potential cues about heading direction are also available. However, in the case of translation with eye rotation, some decomposition of the different flow components is necessary before the global focus of outflow, invariantly represented in the optical flow field, can be recovered from the retinal flow (Regan and Beverley, 1992). Before we can consider whether the response characteristics of MST cells are appropriate for ego-motion representation, we need to examine the effect of manipulating the informational content present in flow stimuli both on navigational performance and MST single unit activity. We can then attempt to correlate how these input variables effect psychophysical performance with how they effect response selectivity in MST. A close relation between the two would suggest a functional link. Computation models of DOH determination will also be reviewed, in order to assess whether the response characteristics of neurons in MSTd are appropriate for carrying out this task.

Psychophysical Studies

Early studies had concluded that subjects are quite poor at determining direction of heading based on optical flow alone, with or without eye rotation. However, in the 1980s several groups working independently determined that heading accuracy under a wide range of conditions is approximately 1-2 degrees (Warren et al, 1988), sufficient for the behaviorally important task of obstacle avoidance, e.g. smashing into trees while running. Similar results were obtained even for the more complex case of curvilinear movement (Warren et al, 1991; Turano et al, 1994). One study (Regan and Beverley, 1982) reported heading judgments as accurate as 0.03 degrees for rectilinear translation based on detecting the maximum of divergence in the flow field, but it is now widely agreed that this probably was a consequence of static positional cues in their displays that are not available under naturally occurring situations. Although many of these studies used evolving flow displays with other cues present besides the instantaneous velocity field, high performance was maintained even when acceleration was removed from the stimuli (Warren et al, 1991). When the global orientation of the velocity vectors was kept intact, but the length of these vectors was randomized (a "direction" field), heading judgments remained accurate, but not when the length of these vectors was left unaltered and the vector orientations were randomized (a "speed" field) (Warren et al, 1991). Interestingly, units in MSTd display a analogous response behavior, showing strong selectivity for direction fields but not for speed fields (Tanaka et al, 1989). Other studies showed that even when considerable noise is added to the stimuli, heading determination remained robust (van den Berg, 1992). In addition, reducing stimulus exposure time had little effect on performance down to 150 msec, much less time than is supplied by the average intervals between saccades. These studies have provided strong support that optical flow is used to determine direction of heading under the limitations imposed under naturally occurring conditions.

Probably the most relevant perceptual data concerning direction of heading determination and the possible necessity of an extra-retinal smooth pursuit signal comes from separate papers by Warren (1988ab; 1990) and Royden, et al (1992). Warren claims human performance is highly accurate for simulated approaches to a horizontal plane, a vertical plane, and a cloud of dots. This was true even when subjects tracked a feature (through smooth pursuit eye movements) that was away from the direction of heading. However, if the rotational component to the retinal flow was added directly to the visual stimulus on the screen, simulating the effects of eye rotation, performance dropped for the approach to a vertical plane, but remained high for the two other conditions. Note that in both the actual and simulated eye rotation sets of conditions, retinal flow was identical. Any performance change had to be a consequence of extra-retinal information about eye movement. Warren concluded that subjects could determine direction of heading from optical flow alone, with or without eye rotation, as long as depth variation was present in the stimulus.

Royden, et al (1992) confirmed many of Warren's findings, but obtained much poorer performance in environments containing depth variation under the condition of forward translation plus simulated tracking eye movements. He concluded that an extraretinal signal is required to determine direction of heading, even when the environment contains variations in depth. A possible way of reconciling these differences is to consider the pursuit speeds used by the two groups of investigators. Warren used pursuit speeds in the range of 0.3 to 1.2 degrees/sec while Royden used speeds as high as 5 degrees per second (van den Berg, 1993). Perhaps the visual system, using information contained in retinal flow alone, can determine direction of heading when the distortions due to rotation are small compared to the signal from observer translation. With greater angular rotational velocities, an extraretinal signal would be required to facilitate this process. Further complicating the story, another group (van den Berg, 1992) concluded that we can determine direction of heading from retinal flow alone as long as there are features very far away from the observer to use as a reference. For these very distant features, retinal flow depends almost entirely on observer rotation, not on translation, remembering that the rotational component of the flow is independent of feature distance. In this study, performance under simulated rotation conditions was poor for the approach to a cloud of dots whose most distant points were relatively near to the subject. It would be interesting to investigate more quantitatively how the signals from translational and rotational flow interacted under more varied experimental conditions. One study of interest (van den Berg, 1992) looked at direction of heading judgments in the presence of noise added to the veridical flow signal. These results showed that an extra-retinal signal becomes more important as the amount of noise in the display increases. We conclude that under optimal conditions direction of heading computations can be carried out by the human visual system based on optical flow alone but that an extraretinal signal is required as conditions deteriorate either because of a lack of depth variation in the stimulus, noisy conditions, or high rotational eye velocities. A somewhat analogous situation exists in the psychophysical literature with regards to localizing visual targets in space. When the extraocular muscles of the observer are partially paralyzed, subjects have a hard time identifying target location in the dark, but have no trouble in the light (reviewed by Matin, 1986).

Cutting (1992) has taken the stance that Gibson's early approach to ego-motion, emphasizing the importance of optical flow, has resulted in misguided attempts to extract this information, seemingly at all costs, from the pattern of retinal flow which is immediately available to the visual system. He argues that no such recovery process is necessary, and that information available in retinal flow is sufficient for recovering direction of heading. He proposes that we use assymetries in the retinal flow field, such as motion parallax, to guide successive eye movements towards the direction of heading. His proposal has several deficiencies. Warren demonstrated that we can accurately recover direction of heading from displays that contain no motion parallax. Additionally, both Warren and Royden conclusively showed that repeated periods of smooth pursuit followed by a saccade to a new environmental feature were not required for accurate navigation.

Models for Determining DOH

Demonstrating that observers can determine direction of heading from retinal flow without an extra-retinal smooth pursuit signal does not provide an algorithm for how this might be done. Knowing the algorithm used by the visual system to recover DOH would be instrumental in determining whether or not MST is an appropriate candidate for this function.

Despite the non-equivalence of the outflow focus and DOH under conditions of eye rotation (with respect to retinal flow), it is common in the psychophysical and computational literature for authors to defend Gibson's approach. The claim is made that under conditions of eye rotation the nervous system must first decompose the flow into translational and rotational components before using this singularity for navigational purposes. What exactly is meant by this decomposition is not at all clear. Neurons respond to stimulus features, such as (potentially) the global focus of expansion. If eye rotation is present, this cue is not a reliable indicator of heading direction. While many of the models discussed below, including those based on differential invariants and differential motion, claim to decompose flow, they are really just extracting a feature from the flow field that, unlike the focus of expansion, is invariant to rotation. If a particular region of the brain is determining direction of heading in this manner, we would expect its neurons to respond to these invariant properties rather than the global focus of expansion. Support for these models cannot be considered as vindicating Gibson's hypothesis, as the flow properties being extracted are not derived from, nor dependent on, the stimulus component he proposed. We will consider one model, which uses an extra-retinal signal in combination with the global focus of expansion, that can legitimately be considered a decomposition scheme consistent with Gibson's proposal.

Some models for determining DOH only have relevance in the artificial intelligence community. For example, discrete models rely on measuring the locations and displacements of a small number of points to set up a system of equations to recover the 6 parameters (3 rotational, 3 translational) that will uniquely describe observer motion. Four to seven points in two successive frames are generally sufficient with this class of model (Warren and Hannon, 1990). These algorithms are obviously non-biological, but are useful in setting lower limits on the amount of information required for these computations. They are notoriously prone to noise as they rely on extremely accurate measurements of displacement and position. Another class of models that has been popular in the computational community is characterized by a least squares search for possible surface layouts and observer motions that are consistent with the observed flow. Although these models are generally resistant to noise and degrade gracefully, they often depend on severe assumptions about surface layout in order to narrow the parameter space searched. Recently, T. Poggio proposed an algorithm that takes advantage of Green's theorem to detect the expansion component of optical flow in a positionally invariant way.

Although this model is robust to noise and its implementation is biologically plausible, he does not offer a way to recover direction of heading once these types of flow fields are detected (Poggio et al, 1991).

A second class of models based on computing the differential invariants (Koenderink and van Doorn, 1981; Koenderink, 1986; Regan and Beverley, 1982) divergence (div), curl, and deformation (def), recognize that these properties of flow fields are invariant with respect to eye rotation. This is because the effect of eye rotation on retinal flow is to add a constant velocity vector on top of every velocity vector representing the optical flow. Spatial derivatives like div, curl, and def, which are dependent only on differences between vectors, are unaffected by such a homogeneous transformation. One of the appeals of this class of model is that information about the structure of the environment is also available from these operations, and many structure from motion algorithms have been shown to use these operators successfully. Unfortunately, for purposes of ego-motion representation this dependence on environmental structure turns out to be a liability. Although these operators are invariant with respect to eye rotation, they do not possess this property with regards to environmental layout. For example, Regan and Beverley (1982) have proposed that we determine direction of heading by detecting the maximum of divergence in the optical flow field, which is not the same as the focus of expansion. Although the maximum of divergence is invariant with respect to gaze direction and eye rotation, it is rarely in the direction of heading. Even if (as they claim) observers are very sensitive to this flow feature, it is not clear how they could use this information to extract information about self motion. In addition to this difficulty, these models are sensitive to noise and require dense flow patterns with smooth depth variations over which to take spatial derivatives. Perceptual experiments discussed above show high levels of performance under noisy conditions, sparse displays, and with environmental layouts such as clouds that preclude the taking of spatial derivatives (Warren et al, 1988). Furthermore, units in MSTd have been shown to be ineffective in extracting these differential invariants from more complex flow patterns (Orban et al, 1992).

Several models which rely on differential image motion to recover DOH have received considerable attention (Rieger et al, 1985; Hildreth, 1992). Differential motion, which includes both edge parallax and motion parallax, occurs during observer translation when two environmental features which occupy nearby locations in the optic array are at different depths. The magnitudes of the velocity vectors associated with these two features depend on their depth, with image points located closer to the observer having a greater retinal velocity for the same eccentricity. If no head or eye rotation is occurring, the orientation of these two vectors is identical, and they point along the direction of heading. However, with a rotational component to the flow, these two vectors are no longer pointed in the same direction, and neither point in the direction the observer is translating. However, if a new vector is created by subtracting these two vectors from one another, this new vector is oriented such that the direction of heading falls somewhere along an extension of its axis. Two such subtraction vectors at different spatial locations uniquely determine the direction of heading by the intersection of their axes of orientation. This procedure can be repeated across the optical array, creating a new positional vector field, whose focus of expansion is aligned with the direction of heading.

Recently, a related model using the affine coordinate system in place of the Cartesian has been proposed (Beusmans, 1993) which also relies on local variations in depth to recover direction of heading. Affine flow is zero for the approach to a vertical plane; it is easy to show that the entire class of differential motion models fail in environments without depth variation. This is consistent with our previous discussion of the ambiguities present in the positional velocity field representing the approach to a vertical plane. Another problem with differential motion algorithms is that they recover only the axis of translation. They do not indicate the translational direction along this axis and therefore whether the observer is receding or approaching from the scene. Information about the relative depth of features in the optic array can provide this information. This class of models is appealing because several groups (Regan et al, 1982; Warren et al, 1988) have shown depth variation in the stimulus improves heading performance when extra-retinal information about eye rotation is absent. As discussed below, to date little sensitivity to differential image motion has been demonstrated for units in MSTd.

An important consideration, under emphasized in the literature, is the fact that we have direct access to optical flow information under a number of situations which leave the eyes stationary relative to their orbits. If these events are frequent enough to effectively guide behavior, further discussion of various decomposition schemes is unnecessary. It has already been mentioned that eye rotation is non-existent when looking in the direction of heading and when fixating on any distant feature. However, such moments of stability also occur for short periods of time immediately following saccades and may even occur briefly during pursuit. If these periods are long enough for the nervous system to detect the focus of expansion and occur frequently enough to provide a behaviorally sufficient sampling rate, we have a very simple solution to our problem. More work needs to be done comparing the frequency and duration of these rotation-free periods to the minimal time required to detect the outflow focus.

MST and direction of heading

One of the strongest indications that MST may be involved in ego-motion analysis is the presence of an extra-retinal smooth pursuit signal that is independent of visual input. All the previously discussed models recovered direction of heading from retinal flow information alone. However, consideration of the psychophysical literature indicates that an extra-retinal eye signal plays an important role in factoring out the flow contribution of eye rotation under non-optimal visual conditions (see above.) From a computation point this task is trivial. Because the effect of eye rotation on the flow field is invariant to environmental structure, an extra-retinal signal can unambiguously represent the component of the flow field that eye movements contribute. By subtracting this representation from that of the retinal flow field, the nervous system is left with a representation of the optical flow field independent of eye movement. This information is then ready to be fed forward to a network which localizes the focus of expansion in this recovered field. At this point, direction of heading can be unambiguously determined, but the coordinate system that this singularity is represented in remains retinal. Its location on the retina changes with eye, head, and trunk movements. Positional information about these body parts will have to be taken into account before the direction of heading can be represented in a coordinate system that is behaviorally useful to guide navigation. We are left with a two step process, one step to factor out head and eye motion and a second to factor out head and eye position. How this latter process is thought to be accomplished in posterior parietal cortex has been investigated by other workers in our lab (Andersen, 1989). Given that MSTd has been shown to respond to the type of flow patterns produced by both observer rotation and translation, and given the existence of an extra-retinal signal for smooth pursuit eye movement present in this region, it is not surprising that MST has been implicated in ego-motion. It has yet to be determined how the extra-retinal signal effects the tuning properties for visual stimuli, an interaction which, if consistent with a mechanism for determining direction of heading, will firmly establish MST as an ego-motion center. These experiments are currently under investigation in our lab. It would also be worthwhile to look for an eye and/or head position signal in MSTd, independent of the smooth pursuit signal, that would facilitate the required coordinate transformation. Given that there is a strong projection from MSTd to posterior parietal cortical area known to integrate such eye and head (and potentially trunk) signals, MST may well not be involved in this latter process.

Psychophysics has shown that spatial integration of motion signals is important in direction of heading determination. Performance remains high under very sparse viewing conditions and large amounts of local noise, as long as the global pattern of the velocity field remains intact (van den Berg, 1992). In addition, the focus of expansion need not be present in the display for accurate heading judgments to be made (Warren, 1992). Subjects even perform well when the direction of translation is off the screen. These considerations rule out a local algorithm where the nervous system recovers DOH by directly detecting the singularity in the optical flow where the velocity signal is zero. Information over a large spatial extent is pooled in making this determination and the global pattern of the velocity field appears important. This is consistent with the response characteristics of MSTd cells. Responses are brisk to random dot stimuli even under very low density conditions and when the focus of expansion falls outside of the receptive field. In addition, MST cells respond well to flow-type stimuli that have the magnitude of their velocity vectors, but not their orientation, randomized. Direction of heading judgments are also robust with these "direction fields" fields (Tanaka et al, 1989). Both MSTd response selectivity and ego-motion judgments fall off in "speed fields" where the speed gradients associated with the flow stimuli are maintained but the orientations of these vectors are randomized. As discussed above, MSTd cells have a preference for isotropic radial stimuli over axial motion pattern fields (Tanaka et al, 1989). As movement through the environment produces isotropic patterns and not axial ones, this selectivity again supports a role in MST processing optical flow, although not specifically for the function of determining direction of heading.

As mentioned above, 40 percent of MST cells tuned to disparity reverse their directional selectivity for linear motion with changes in relative depth (Roy et al, 1992). The investigators who documented this finding proposed that cells with this tuning behavior would be ideal candidates for registering self-motion. They pointed out that when translating forward and fixating a feature away from the direction of heading, objects in the line of site behind the fixation point move opposite to the direction of pursuit, while those in front of the fixation point move in the same direction, but at a greater velocity. Cells which reverse their tuning with changes in disparity would be ideally constructed to register this stimulus. However, how this information could then help recover direction of heading is not clear. It is also not clear if the range of distances over which disparity is known to be detected is sufficient to be helpful in this process, but it remains a tantalizing possibility. Finally, on more theoretically grounds, Lisberger, et al (1978) argue that cells with large receptive fields, and not retinotopically organized ones such as those found in MT, are required for a "sensory-motor interface." An ego-motion center would appear to belong to this class of detector, and the large receptive fields of MSTd cells would support its candidacy.

Most groups, including ours, have shown a significantly higher percentage of MST cells preferring expansion patterns than any of the other radial motion types (Graziano et al, 1992; Tanaka, et al 1989). One investigation by Orban's group (Lagae et al, 1994) found more rotation cells than those tuned for expansion, but this was probably a sampling error. The apparent disproportionately large number of cells tuned to expansion is consistent with the frequency that this type of optical flow pattern occurs under natural conditions. Backward motion is a relatively rare occurrence, and we would expect that a region processing ego-motion information would be underrepresented with cells tuned to this motion type. It can be argued, however, that expansion may be environmentally more relevant as a stimulus, independent of its role in determining direction of heading. For example, as discussed below, looming objects generally pose a much more impending threat that receding ones. The over representation of expansion cells may reflect the relevancy of this type of motion pattern for a completely different process. It is important to re-emphasize that to claim an area is processing optical flow is not to say that the area is recovering direction of heading. Many other functions, which we will come to shortly, potentially have flow stimuli as their inputs.

A strong case can be made for MST having little role in the computations that recover direction of heading from optical flow. A great deal has been made about the positional invariance of these cells with respect to their preferred tuning (Orban et al, 1992; Saito et al, 1986; Tanaka et al 1989; Duffy et al, 1991). Positional invariance does not come cheaply in terms of creating this type of response selectivity with inputs from unidirectional motion detectors such as those found in MT. Although the amplitude of the response to the preferred stimulus pattern is affected by the location of the stimulus center in the receptive field, the response profile is fairly flat with respect to this placement. For a region of the brain supposedly involved in recovering direction of heading, we would expect that the location of the center of flow to have a large effect on the way these units respond. Although this does not rule out a coarse coding scheme to represent this information in a distributed way, clearly MSTd cells in isolation are not coding direction of heading.

An interesting possibility exists that the large degree of positional invariance observed in MSTd has to do with the choice of stimuli. In every study that examined this feature of MSTd cells, stimuli without any depth variation (and therefore without differential motion) were used. Psychophysical studies (discussed above) have shown that direction of heading judgments toward surfaces such as vertical planes, which have no depth variation, are degenerate cases and that performance under such situations is often impaired. Maybe if MST cells were presented with stimuli such as clouds of dots or horizontal (ground) planes that do contain differential motion, less positional invariance would be observed. Our lab is currently investigating this interesting possibility.

In many species, the selectivities of neurons in regions of the brain thought to be involved in analyzing optical flow are often restricted to a limited number of cardinal directions. For example, in the rabbit accessory optical system (AOS), the preferred tuning of these unidirectional motion detecting units is restricted to the three axes which correspond to the orientations of the semicircular canals of the inner ear. Although the orientation of flow field lines can vary continuously, depending on the movement of the rabbit, the limited selectivity range of units detecting this motion effectively decomposes the flow into three separate channels. The same approach is taken in several other species. Single unit recording in the fly lobula plate has demonstrated that cells tuned to global motion patterns respond preferentially to horizontal and vertical motion, but not to intermediate directions (Egelhaaf et al, 1988). Three distinct channels tuned to linear motion corresponding to the orientation of the semicircular canals have been reported in the flocculus of the pigeon (Wylie et al, 1993ab; 1991).

Because the eyes of many organisms are oriented laterally, the flow patterns projected onto their retinas consist of parallel fields of velocity vectors, rather than the radial type patterns observed by animals which have their eyes situated more anteriorly. It might be expected that higher primates would take a similar decompositional approach, with a channel specificity appropriate for the type of flow patterns they are faced with. Computational models of ego-motion which use spatial derivatives to decompose the scene into deformation, divergence, and curl (rotation) apply this approach. A simple implementation of such an algorithm might postulate units selective for these components of the flow. If the divergence component of the flow could be extracted from more complex stimuli, MST cells would directly represent an aspect of flow independent of head and eye movement. Graziano, et al (1994) ruled out such a channel hypothesis for MSTd by finding cells selective to spirals, which contain components of both curl and divergence. In addition, several groups have demonstrated that single cells in MST are selective for both radial type stimuli and linear motion patterns, arguing that a decomposition is not being performed along this dimension either. Furthermore Orban, et al (1992) reported that MST cells could not extract the elementary flow components (EFCs) of expansion, curl, and deformation from more complicated stimuli that contained a mixture of these elements. For example, when a cell tuned for expansion had a field of rotational flow added transparently on top of this stimulus, the cell's response became detuned. Although these findings cannot rule out MST having a role in direction of heading determination, they do rule out a potentially economical way of decomposing the optical flow that has apparently been exploited by other species. The fact that Old World monkeys do not use such a strategy may simply reflect the greater computational difficulty of decomposing the more complex flow fields that they are exposed to as a consequence of anterior eye placement and capacity for smooth pursuit eye movements.

Erickson, et al (1991) presented findings in MSTd which are also problematic for assigning this area an ego-motion registering role. They showed that units in this region responded to patterns of motion moving in the environment, but not to the equivalent retinal motion produced by making a smooth pursuit eye movement over the pattern. This seems to suggest that MST cells are analyzing environmental motion and not movement of the observer through the environment. They found that the same was true for MSTl, but not for V4 or MT, which responded identically, regardless of how the retinal motion was produced. If MST is specialized for ego-motion representation, we would have expected the opposite result. It would be interesting to see if the same distinction holds for the rotational and expansional components of flow. If translating the monkey through space fails to activate these units, despite presenting similar motion patterns on the retina as those effective in previous single unit recording studies, it would be difficult to assign an ego-motion role to this cortical area.

Another puzzling property of MSTd cells is that although their receptive fields generally include the fovea, the center of these fields is often eccentrically placed. This is significant, considering that the focus of expansion of the retinal flow during translation is in the direction of gaze, and not eccentrically positioned. The only time this is not the case is when the eyes are stationary in their orbits. If MST cells are involved in analyzing optical flow, it might be expected that the centers of their receptive fields should be located at, or very close to the fovea. The suggestion, made above, that perhaps judgments of direction of heading are only made during the short intervals when angular orbital velocity is zero, may explain this eccentric placement. Also important to reconcile is the absence of a vestibular signal in these cells. Given the importance of vestibular information in guiding locomotion, it is hard to understand why MST is not taking advantage of this information. Maybe the function of MST is limited to processing ego-motion information available from visual cues and feeds this information forward to a cortical region where different modalities related to DOH are integrated.

Although it has been paid less attention in the literature, something should be said with regards to MST's potential for representing information about ego-speed. MSTd units generally start responding significantly at stimulus speeds of approximately 1 degree/sec and gradually increase their response amplitudes with stimulus speeds up to 20 degrees/sec. Rarely is band-pass speed tuning observed, where response drops off after a certain velocity is exceeded. This broad tuning (Tanaka and Saito, 1989; Saito et al, 1986) would make it difficult for this region to accurately encode speed, although it cannot be ruled out that the information is present in a distributed manner.

Recently, a neural network model was trained to recover direction of heading from optical flow, even in the presence of eye rotations (Lappe and Rauschecker, 1993). The network received no extra-retinal signal and was able to extract information about observer motion based on retinal flow alone. This model had only two layers. The input layout was inspired by the physiology of area MT; the units had relatively small receptive fields tuned to two dimensional unidirectional motion. After training, the output layer took on many of the receptive field properties as cells in MSTd. Receptive fields were large and tuned to all combinations of radial motion as well as translational motion. Interestingly, all of these units were of the triple component type. Each unit responded, to some extent, to translation, expansion/contraction, and rotation. Unlike area MST, no single or double component units arose as a result of the training process. Also, although these triple component units exhibited some degree of positional invariance in their tuning, there were boundaries within their receptive fields where their motion pattern selectivity reversed, e.g. the units responded to expansion in one part of their receptive fields but reversed their selectivity to that of contraction in another part. Their model depends on these reversals to encode heading direction. Although some cells in MST exhibit this property, we have found that they are the rare exception (unpublished observation). Interestingly, under noisy conditions and as the depth variation in the input stimulus decreased, performance of the model deteriorated, consistent with human performance when extra-retinal information is not available. They observed that their output units became less positionally invariant, with respect to amplitude of response, as the depth of the input stimulus was reduced. As mentioned above, our lab is currently investigating the effect of varying depth (as reflected in differential image motion) on the response properties of cells in MSTd.

In light of the objections raised above with regards to the suitability of MST in computing direction of heading, are there regions of the brain that are better candidates for this function? In several species, the accessory optic system (AOS) has been implicated in this role. Unfortunately, the physiology of this region has not received much attention in primates. This research is particularly needed in the primate because animals in which this system has been studied in detail are exposed to very different flow patterns because of lateral eye placement.

The AOS has been extensively studied in birds (Frost et al, 1990). This literature contrasts the ego-motion role of the AOS with that of the tectofugal system, which is thought to be involved in the analysis of object motion in the environment. Cells of this latter neural system are directionally selective for the movement of small stimuli. These cells have a double opponent organization to their receptive fields, meaning that they prefer opposite directions of motion in their centers and surrounds. Surround motion in the same direction as that of preferred motion in the center inhibits the response. Therefore, wide-field motions, such as those present in flow stimuli, activate these cells very poorly. The AOS, in contrast, responds poorly to small stimuli and prefers large, textured motion patterns. Furthermore, cells in one of the component nuclei of this system, the nBOR, have binocular receptive fields and respond best to opposite directions of motion in the two eyes (Wylie et al, 1993; Wylie and Frost, 1993); they therefore respond best to optical flow created while moving forward in the environment. These cells have no inhibitory surround. Similar cells have also been studied in the AOS of the turtle (Rosenberg and Ariel, 1990). The AOS has also been studied extensively in rabbits (for a review see Simpson et al, 1988). The rabbit AOS has three components, the medial, lateral, and dorsal terminal nuclei. The dorsal terminal nucleus (DTN) prefers unidirectional motion along the horizontal axis while cells that comprise the other two areas prefer near-vertical motion. Cells in all three nuclei have large, contralateral receptive fields and prefer large, textured stimuli, consistent with their proposed role of analyzing flow. In all regions of the AOS studied to date, these directional cells prefer speeds in the range of 0.5 degrees/sec, about the speed of environmentally induced flow stimuli.

A direct projection from MST to the AOS has been demonstrated in the macaque (Maioli et al, 1989). Possibly, these two areas work in conjunction to encode heading direction. A particularly appealing theory suggests that MST evolved in animals to deal with the type of flow patterns that arose with anteriorly placed orbits and vergence eye movements. More investigation of the primate AOS is clearly called for. Unfortunately, these studies are difficult because of the relatively small size of these nuclei in the monkey.

Many of the cells in the dorsolateral pons of the macaque have receptive field properties similar to those in MT and MST, from which they get projections. Their receptive fields are generally broad and are directionally tuned. Some units prefer large, textured stimuli, while others prefer the translational movement of small objects (Suzuki et al, 1990). To our knowledge, these cells have not been tested for selectivity to radial stimuli like expansion and rotation. Cells from the dorsolateral pons then project to cerebellar areas known to be important in such eye movement related behaviors as the optokinetic reflex. Given their known involvement with smooth pursuit and the selectivity of at least a population of their neurons to wide-field stimuli, this region should not be ruled out from being involved with visually guided navigation.

Both the AOS and the pontine systems project to the cerebellum. In a series of careful studies conducted on the Purkinje cells of the pigeon (Wylie et al, 1993; Wylie and Frost, 1991, 1993), neurons were found that preferentially responded to the types of flow produced by descent, ascent, roll, and yaw movements of the bird. These authors propose that, in the pigeon, both the AOS and the cerebellum play a role in analyzing flow information and that they represent different stages along the flow processing pathway. Most AOS cells are monocular and it is not until the cerebellum that motion signals are integrated from the two eyes. Studies in other animals have also found cerebellar neurons selective for flow-like stimuli in the frog (Ansorge and Gusser-Cornhis, 1977), rabbit (Ghelarducci et al, 1975), and monkey (Waespe and Henn, 1981). Clearly a complex network of subcortical structures is in place for analyzing optical flow, and any proposed role for MST cells should be considered in this context.

What about other cortical areas that could potentially be involved in processing self motion? Albright (1989) reported that MT cells representing the periphery of the visual field have more cells tuned to motion along a centrifugal (away from the fovea) orientation than any other direction. As centrifugal motion is induced by forward locomotion and this is the most frequently observed type of observer translation, this centrifugal bias is consistent with MT analyzing motion from this source. A similar organization in tuning bias is observed in cells within the cat suprasylvian cortex (Clare Bishop area and the PLLS) (Nawrot and Blake, 1993). Although it may be unlikely that MT can extract direction of heading information in isolation, it may be performing some initial processing of the flow signal for this purpose.

Other candidate regions for processing self motion are areas of parietal cortex. There are reports in the literature on "parietal visual neurons" possessing directional selectivity that is opponently organized along meridians. These cells respond to features moving either towards or away from the fovea and investigators (Steinmetz et al, 1987) have suggested that such units play a role in analyzing optical flow. Although it cannot be determined from this study whether cells in this region were receiving projections from MST, it is not unlikely given the area from which they were recording. Neurons from area 7a, which is known to receive a projection from MST, have been reported with a similar selectivity for motion towards or away from the fixation point (Motter and Mountcastle, 1981). Like MST, neurons in this area have been shown to respond, in a positionally invariant way, to rotational stimuli (Sakata et al, 1986). In yet another region of cortex, cells in the rostral polysensory area have been shown to be tuned to both expansion and contraction (Bruce et al, 1981; Hikosaka et al, 1988). The convergence of information from several sensory modalities in this region may provide a forum for the integration of diverse cues about ego-motion. Finally, vestibular cortex, which is thought to play a role in the sensation of vection, may play a vital role in understanding the way we navigate through the environment.

In conclusion, given that many cortical and subcortical regions possess response properties appropriate for the analysis of optical flow, we should not assume that this role is exclusively subserved by MST. It is conceivable that MST is not involved at all, or only plays a part in recovering DOH from optical flow. The processing of optical flow information for purposes of navigation arose very early phylogenetically (the fly can perform these determinations) and it is quite possible that cortical centers, if involved at all, may well have evolved on top of, and supplement, a pre-existing more primitive system. The known connection between MST and the AOS supports this notion.

Vection

Vection is another perceptual phenomenon that has optical flow patterns as its input. Vection is the sensation of moving relative to the environment and although classically thought of as a visual/vestibular interaction, can also be induced by changes in the acoustic surround and by tactile motion (Andersen, 1986). Vection is believed to be intimately related to motion sickness, which is thought to be induced by a mismatch of vestibular and visual information. It is likely that the determination of direction of heading and the experience of vection are related. For example, psychophysical studies have shown that differential motion, arising from depth variations in the inducing stimulus, is important in some types of vection (Telford et al, 1992), as well as for direction of heading perception. Similarly, the maximum of divergence of the positional vector field representing flow is not thought to be important for either inducing vection or recovering heading direction (van Esten et al, 1988). Also, the minimum velocity conditions required to induce vection are similar to those for detecting flow and motion in general, arguing against two completely separate motion processing systems (Andersen, 1986). However, there are a number of studies which show an independence between the two systems. For example, the corrections made for ego-motion in order to accurately represent the motion of objects in the environment occurs independently of the sensation of vection (Brenner, 1993).

Linear vection is the sensation of translating through the environment, circular vection is the sensation of rotating about the vertical axis (orthogonal to the ground), roll vection is the sensation of rotating about the line of sight, and pitch vection is the sensation of rotating around the horizontal axis. Given that the patterns of visual stimulation that will trigger all these types of vection have been shown to effectively drive MST cells, (they are the same set of stimuli relevant for direction of heading determination), it is not unreasonable to suggest that MST has a role in vection perception. Although early reports suggested that vection was more effectively induced by stimulation of the visual periphery, later studies showed the sufficiency of presenting the inducing stimuli foveally (Andersen and Braunstein, 1985; van Esten et al, 1988; Telford et al, 1992). This is consistent with the observation that a majority of MSTd neurons include the fovea within their receptive fields. Additionally, the relative scarcity of MST cells tuned to rotation compared with translation and expansion may explain why the perception of roll vection is generally so weak. Furthermore, whether a flow stimulus is perceived as being part of the foreground or background is relevant to its effectiveness in inducing vection (Telford et al, 1992). MST cells broadly tuned to disparity might be employed to help make this distinction if this region was involved in this processing.

Despite these arguments, there is a good deal of evidence that activity in MSTd does not correlate well with the perception of vection. The visual stimuli used in our lab, although at times producing the impression of an object moving towards the observer, never induced a sense of vection in humans. Despite this, these stimuli drove the units we studied in MSTd quite vigorously. Although it is impossible to determine what the monkey was experiencing during recording sessions, in ignorance we will assume that their perceptions are similar to our own. Additionally, psychophysical experiments have demonstrated that the induction of vection using centrally placed stimuli requires that the display contain depth cues provided by differential motion (i.e. motion parallax) (Telford et al, 1992). We have already reviewed MST's apparent insensitivity to such cues. Some studies have shown that, besides differential motion, vection is sensitive to deformation in the inducing stimulus (van Asten et al, 1988). This observation again argues against MST, which has been shown to have very few units specifically sensitive to this type of motion (Lagae et al, 1994).

The ratio of cells tuned to expansion versus contraction in MSTd is quite high. Because of this, if MST were encoding linear vection, we would expect that latencies for inducing forward vection might be shorter than for inducing backward vection. The opposite has actually been observed (Andersen, 1986). Finally, the dynamic response profiles of MST cells and vection onset are not well matched. The time elapsing between stimulus initiation and vection is generally about 2 seconds (Fluckiger and Baumberger, 1988), a much longer period of time than the approximately 80-120 msec delay between stimulus onset and MST response.

It has been proposed that motion information is processed independently by two separate pathways, a geniculo-striate pathway and a retino-tectal pathway. This latter system has been thought to be involved in "blind sight" in cases where, despite destruction of striate cortex, some visually induced behaviors can still be elicited. Some workers have raised the possibility that this alternative system carries the inputs for vection perception. However, a study examining human stroke victims (Straube and Brandt, 1987) with lesions in primary visual cortex showed that these patients did not experience vection when presented with stimuli in the blind portion of their visual fields. As the alternative retino-tectal pathway, consisting of an ascending projection to the pulvinar nucleus and then extrastriate cortex, was intact, it was concluded that this latter system was not involved in transferring visual information responsible for vection. A second group of patients, with lesions to vestibular cortex, also did not experience vection, despite having an intact primary visual cortex and normal perception for object motion.

The balance of evidence reviewed above suggests that the same pathway is involved in both processing direction of heading information and the sensation of vection. However, vection is likely realized much later in this processing stream than either object motion or ego-motion, possibly in vestibular cortex. It is unlikely that MST serves as more than an early processing center for this perceptual phenomenon.

Postural Stability

Postural stability is clearly related conceptually to both vection and ego-motion and may serve as a sort of link between these two processes. Both vection and ego-motion are both clearly sensory in nature while postural stability can be thought of as a subconscious motor function that receives input from one or both of these systems. Obviously other sensory systems, particularly the vestibular, play a role in maintaining postural stability; we will limit this discussion to those aspects of postural stability influenced by optical flow stimuli and therefore potentially mediated by MST. Although we are restricting our scope to visual processing, we include under the heading of postural stability any non-voluntary action, including the optomotor response, affected by changes in the optical array.

Optical flow has shown to have many motor consequences, including head bobbing in pigeons, locomotion in lobsters, and compensatory postural adjustments in humans (Pailhous et al, 1990; van Asten et al, 1988; Fluckiger and Baumberger, 1988; Gielen and van Asten, 1990). It is thought that many of these behaviors are "hard-wired" into the nervous system and involve rather low level processing mechanisms. Postural changes as extreme as falling have been elicited in infants on the basis of visual motion stimuli alone. With age, these types of responses are generally attenuated, suggesting a higher order modulation of these systems through experience.

Is MST an appropriate candidate for the low-level sensory input processor to these systems? Psychophysical studies on humans have been ambiguous in their conclusions about the response characteristics of the perceptual part of this network. Some groups report that postural responses to optical flow are invariant with respect to environmental structure (van Asten et al, 1988), consistent with the response profile of MSTd cells (see Chapter 2). However other groups (Fluckiger and Baumberger, 1988) have found the opposite. Consistent with the relative lack of speed tuning of this cortical area, postural responses are invariant with respect to velocity (van Asten et al, 1988). However, this same paper reported that motor responses were dependent on the texture of the inducing stimulus, in conflict with unit selectivity in MST. In terms of an association between ego-motion and vection, yet another investigation (Gielen and van Asten, 1990) showed a dependency of vection on fixation direction, a finding inconsistent with similar experiments on ego-motion perception. In addition, optical flow can influence subconscious motor behavior even with an interruption of the geniculo-striate pathway (Mestre et al, 1992), perhaps through the colliculus-pulvinar-parietal pathway.

It is not clear from these studies the extent of the interaction between the systems processing vection and ego-motion or whether either function is processed in MST. As the data is equivocal, it is hard to make any firm conclusions with regards to the role MST plays in inducing postural changes. However, it is reasonable to assume that several different motor systems are involved because the behavioral consequences of presenting flow stimuli are so varied, e.g. from inducing smooth pursuit eye movements with the optokinetic response to changing stride length during walking. Given MST's selectivity for optical flow type stimuli, we tentatively propose that MST is part of the visual pathway involved in processing the sensory component of these postural responses, but that other areas are likely to be important. How close MSTd lies to the sensory-motor interface of these systems is difficult to say. It is also quite possible that several parallel pathways are involved and that the role MST plays may be only facilitory, and not essential.

Retinal Slip

Another possible role that MST plays, suggested by its anatomy and physiology, is of detecting "multi-dimensional retinal slip." Retinal slip has long been recognized to help drive smooth pursuit. When smooth pursuit motor activity is insufficient to keep a feature perfectly foveated, the projection of that feature moves on the retina. This causes cells which are sensitive to linear motion (such as MT cells and many MST cells) to respond and presumably feed a signal into the smooth pursuit system to adjust motor activity in the proper direction to correct for the slip. As an observer moves through the environment and tracks a feature, a radial pattern (generally a spiral) of flow gets projected onto the retina. As long as the feature remains perfectly foveated, the center of outflow for this radial pattern is in the direction of gaze. But if pursuit activity is not accurate in keeping this feature foveated, the spiral retinal flow pattern will "slip" on the retina. Given that MST is sensitive to radial flow patterns, this area may detect this slip and feed an appropriate error signal into the smooth pursuit control system. However, this hypothesis suffers from at least one of the same weaknesses as the ego-motion proposal: cells in MST have considerable positional invariance in their response characteristics, a seeming liability for both of these functions. Despite this, some coarse coding scheme could still accurately represent this information, albeit in a distributed fashion.

If MST is involved in detecting "multi-dimensional" slip, it would explain why the majority of MST receptive fields have the peak in their response sensitivity away from the fovea. As the center of the radial flow pattern moves away from the fovea, it drives these eccentrically placed receptive fields more vigorously. This increased activity could then be fed into the smooth pursuit loop as an error signal. A potential difficulty with this theory is the recent finding that lesions to MSTl, and not MSTd, effect smooth pursuit. However, these studies took place under conditions where the monkey was otherwise stationary while tracking a moving object in the environment. The type of retinal slip registered under these conditions is of a small object being displaced on the retina. Our proposal involves the slip of radial motion patterns across the retina and is most relevant during observer translational movement. This is the first time, to our knowledge, that such a possibility has been suggested. Although, at present, it is a tentative idea at best, it deserves future consideration. Certainly involvement with this activity would not preclude MST from contributing to other types of visual information processing.

Object Motion vs. Flow Segmentation

The heavy backward projection from MST to MT makes it possible that MST cells are responsible for the receptive field surrounds of MT neurons. Given the wide-field selectivity of a large percentage of cells in MSTd, it is likely that only MT cells without inhibitory surrounds are contributing to their receptive fields. An interesting network arises out of this potential specificity of connections: MT cells without surrounds are sending connections forward to MST to create units selective for large-field motion. These units then feed backwards to produce the receptive field surrounds of a second population of MT cells.

This network would be ideally constructed to separately parse motion of objects and motion produced as a consequence of observer movement. To imagine how this would occur consider the following situation: A monkey is moving in a straight line through the forest. For the sake of simplicity, assume that he is looking straight ahead and that his eyes are fixed in his orbits, and his head is immobile relative to his shoulders. An expanding pattern of flow is evolving on his retina due to his translation in the environment. His surroundings, in a spatial (observer independent) coordinate system, are stationary except for a cheetah that is crossing the monkey's path. How can the monkey disentangle which retinal motion is a consequence of his own movement and which is a consequence of the cheetah's movement? Equivalently: How can the monkey distinguish proprioperceptive from exteroperceptive components of retinal flow? From a computation point of view, such a system might take advantage of the fact that the types of retinal flow patterns that occur as a consequence of self motion through a rigid environment are fairly constrained. The recognition that these motion patterns are fairly stereotypical is what originally motivated many workers in MST to suggest that this region is analyzing flow: the selectivities of units in this area are well matched with these constrained arrangements of the velocity field.

If an object is moving in the environment, most likely part of the optical flow pattern provided to an observer detecting this event will violate a rigidity constraint; i.e. a pattern of motion will occur that is consistent with no possible combination of observer motion and rigid environmental layout. In general, the majority of the velocity field is compatible with having a single proprio-specific origin (i.e., arising as a consequence of observer motion), while small image patches are inconsistent with this source. These local violations of rigidity correspond to motion of objects in the environment.

Under some situations, the majority of retinal flow reflects environmental movement rather than observer translation. This may not pose a problem as long as the flow from this object motion is not itself consistent with flow produced from a single set of observer motion parameters through an equivalent stationary environment. This is generally the case when multiple objects are moving independently. For example, if one were to be standing amongst a flock of birds which took off in flight, a considerable amount of motion would be presented to the observer. However, because no possible combination of observer translation and rotation could simultaneously be consistent with producing all of this varied motion, the nervous system (at least in theory) should not be fooled into assigning this flow a proprio-specific source. In contrast, looking out the window of a stationary train while another train passes stimulates large portions of the retina with motion that is consistent with observer translation, despite being produced by the movement of an external object. Due to a failure to veridically parse flow information into extero-specific and proprio-specific components, this situation gives rise to the well known illusion of the observer's train moving.

Designing an ego-motion system from scratch, we might build into it the assumption that the largest part of the optical flow field consistent with the rigid translation and rotation of an observer will be considered as observer induced and be used in direction of heading computations. Through an unspecified process of spatial integration, inconsistencies in the flow field (do to object motion) could be smoothed over to obtain a complete representation of the flow field produced by observer movement. Because areas of optical flow containing object motion generally have greater discontinuities in their velocity vector fields than regions where the environment is rigid, signal smoothing preserves more information contributed by observer translation than object motion. This process is important for the perception of object motion as well, because the contribution to the flow from objects moving across the visual scene is affected by their own motion through space as well as the motion of the observer. To recover the former component, observer induced flow needs to be accurately represented in order to subtract its contribution.

MST is an ideal candidate for this spatial averaging process. As discussed above, these cells respond to the global motion pattern present in the display and tend to ignore local inconsistencies in these patterns. Once motion information reaches this region, these local fluctuations are smoothed over and cannot be recovered. Under this model, MSTd is, in essence, recovering what the flow pattern would have looked like had there been no object motion (like the cheetah in the example above) in the scene. In area V1, with their units' smaller receptive fields and lesser degree of spatial averaging, the details of the flow field are preserved. In a simple implementation of this model, MT would be where the representations in MST and V1 converged, with the units in MT sensitive to differences in the two representations. V1 would contribute to the activity in the receptive field centers while MST would provide the input for the surround. In this conception, MT would be "real motion" cells which represented that portion of the flow contributed by objects moving in the environment.

A problem with this proposal is that MT cells with surrounds do not quite have the required behavior. Although motion in the surround in the same direction as that in the center of the receptive field does suppress these cells' responses (Born and Tootell, 1992), a stationary stimulus in the receptive field center will not drive the unit, regardless of the motion presented to the surround. In more jargony terms, what is needed for such a subtraction to occur is a "double opponent" type organization to the receptive field. Such cells have been found in MSTl (Tanaka, et al 1993). They respond poorly to large, moving, textured stimuli, but well when a stationary object is presented on top of this field. These studies were done with homogeneous fields of linear translation (unidirectional motion). It would be worthwhile to try these same studies using radial stimuli for the background.

It may well be that such a clear-cut segmentation of various flow field representations into discrete regions will not be possible and that only by considering the dynamic interactions of these regions operating together will a full understanding of this process be understood. However, we believe this is a reasonable framework for considering the issues involved.

To sum up, a possible function of area MST is to represent a smoothed flow field so that it can be ignored, to allow for detection of moving objects in the environment. MSTd's spatial integration characteristics and the specificity for the type of large field motion patterns that are produced by observer movement are essential for this task. The second essential component of this model is the necessity for some type of subtraction process whereby the flow contributed by object motion is extracted. Cells possessing "double opponent" selectivity would be ideal for this role.

Reference Copy

An extension of the idea presented in the previous section of "representing the flow so as to ignore it" comes from a proposal in the psychophysical literature in a study of the Filehne illusion (Wertheim, 1987). This well known perceptual effect is as follows: When a subject makes a smooth pursuit eye movement in darkness a small, stationary object presented away from the fovea appears to move in the opposite direction as pursuit. The traditional explanation for this illusion is that an efference copy, extra-retinal signal for pursuit eye movement, feeds into a perceptual correction system that has insufficient gain to compensate for the fact that the environmentally stationary target is moving on the retina. The observation that this effect does not occur in the light suggested to the investigators that visual information, in the form of flow induced by eye rotation, feeds into this same correction system and augments the extra-retinal signal. This combined sensory and motor input they call a "reference signal." Supporting this proposal, they determined that the extent a visual reference could attenuate, and even reverse, the Filehne illusion positively correlated with the optokinetic potential of the stimulus. Consequently, a background pattern composed of low spatial frequency luminance components (high optokinetic potential) resulted in the target appearing to move in the same direction as smooth pursuit, creating a "reverse Filehne" illusion. They concluded that "Retinal images not only generate retinal signals, they can also affect or generate reference signals and thus contribute to, or even determine, their own perceptual interpretations."

The proposal of the previous section is largely an extension of this same idea, to include radial flow patterns produced during observer translation. The fact that MSTd contains extra-retinal smooth pursuit activity adds further weight to the hypothesis that MST contains such a "reference signal." There is good evidence that this self/environment distinction is a relatively low level one and that it is hard wired into the human perceptual system. 16 week old infants can distinguish between retinal motion produced by their own movements in the world and the movement of things in their surroundings (Kellman et al, 1987). Other areas of cortex may be involved in similar processes. For example, single unit recording studies in macaque polysensory cortex have located cells which fire in response to unexpected motion in the visual scene but ignore the self-initiated movement of the monkey's arm into the receptive field (Hietanen and

Perrett, 1993). Apparently, an efference copy for limb movement is being sent to this cortical region and could be playing a similar role as the extra-retinal eye signal present in MSTd.

Structure From Motion

Optical flow is important in recovering information about the static structure of the environment. Because MST is selective for flow-like stimuli, this region may be important in this context.

Moving through the environment provides us with a tremendous amount of information about its structure. For example, moving ones head from side to side when one eye is closed will help recover an impression of depth that normally does not exist under monocular conditions. If instead of moving the head back and forth, an eye is rotated instead, no such effect is achieved. Why? With a pure rotation, every feature in the stationary environment moves over the retina with the same velocity (with regards to both direction and amplitude). With lateral motion, although the directions of all the velocity vectors making up the optical flow field are identical, their magnitude is determined by each feature's distance from the observation point. Nearby objects move over the retina with relatively high velocities; objects at infinity do not move at all. For more complex translations, the situation becomes more difficult to visualize. As long as there is no rotational component to observer movement, the depth of a feature plays no role in determining the direction of its motion over the retina, but the speed of this motion is affected by distance from the observer. If rotation is present as well as translation, both speed and direction depend on the movement of the observer and the distance of the feature. The aspects of the flow that are dependent on the depth structure of the environment constitute motion parallax. A related phenomenon occurs when objects move relative to the stationary observer. If these objects contain depth, the displacement of each feature making up the retinal projection of that object depends not just on its trajectory in space, but on its distance from the observer. Features of the object further away from the observer move more slowly across the retina than ones nearby. Rather than giving the impression of a distorting shape, the nervous system somehow compensates for this and each feature appears to move with the same velocity. Like motion parallax, this differential motion provides information about the depth structure of the objects that are moving.

The capacity of the nervous system to extract relative depth information from the optical flow (Nakayama, 1985) is a phenomenon called structure from motion (SFM). We should emphasize that only relative depth information is available from optical flow, not absolute distances. To recover distance, relative speed between observer and objects is required. Along with the other roles that MST may be playing with regards to optical flow, we should consider the possibility that it is involved in SFM processing as well.

This possibility is unlikely. Although lesions of the STS in the region of MST have been shown to impair tasks that involve structure from motion (SFM), these lesions have damaged MT as well (Regan et al, 1992). The large receptive fields of units in MST are inappropriate for both the analysis of local differences in flow and the accurate spatial representation required for SFM computations. While the presence of cells tuned for the differential invariants of divergence and curl are well documented, Orban's group (Lagae et al, 1994) has noted that the relative scarcity of units tuned for sheer and deformation argues against a structure from motion capacity. More direct evidence against MSTd playing a role in SFM comes from single unit recording studies that have revealed a lack of selectivity to the depth structure in the stimulus. Displays of approaching clouds and approaching planes drive MST cells equally well and no more subtle differences in response selectivity to these different conditions has been reported so far. Finally, models which propose an ego-motion role for MST will be hard pressed to integrate SFM computations into this same region. Direction of heading determination, to be robust, must be insensitive to environmental structure, an observation which demonstrates the fundamental incompatibility of the two functions. It is much more likely that SFM information is processed in area MT, whose center-surround organization provides a reasonable mechanism for retinotopically representing differential motion in the optic flow (Masud et al, 1989).

Time To Collision

Yet another important piece of information, time to collision (TTC) is buried within the optical flow. This is somewhat surprising since neither information about observer translation speed nor absolute distances to objects can be determined based on flow alone (Nakayama, 1985). Behavioral studies have indicated that subjects judge TTC on the basis of the parameter tau, which is the ratio of the angular extent of target size over the rate of change of its angular extent (Cavallo, 1988). The accurate measurement of tau has been shown to be important in behaviors such as catching a baseball. In a study that tested subjects in a paradigm that involved driving a vehicle, accuracy of TTC estimation increased with visual field size, binocular vision, higher speeds and driving experience (Cavallo, 1988). With regards to the neural network recovering this parameter, theoretical arguments (Regan and Hamstra, 1993) argue for a system sensitive to isotropic expansion, but not other forms of expansion. To an approximation, MST cells have this selectivity. As reviewed above, MST neurons respond poorly to axial expansions and increase their selectivity as the number of unidirectional motion directions defining the motion patterns increases (Tanaka et al, 1989).

Psychophysical experiments have dissociated the perception of rate of expansion from TTC estimation (Regan and Hamstra, 1993). These studies determined that TTC thresholds are much higher (by two orders of magnitude) than thresholds for detecting rate of expansion. Unlike TTC thresholds, the ability to detect rate of expansion obeys a sort of Weber's law, with sensitivity to fractional differences in expansion rate being approximately fixed across all speed ranges. It is not clear from the available data which of these two stimulus aspects MST cells are more related to, but this would not be difficult to test.

Time to collision neurons have been identified in the nucleus rotundus of the pigeon (Wang and Frost, 1992). Besides being selective for objects on a collision course with the bird's head, the maximum firing rate of these cells always occurs a certain fixed time from collision, independent of the angular extent of the object. This is not true for cells in MST, where the firing rate is affected by the angular extent of the object. Finally, unlike cells in MSTd, nucleus rotondus cells gave poor responses to large stimuli occupying a substantial portion of the visual field. Although we can not rule out the possibility that MSTd is encoding TTC in a less specific way than the collision cells in the pigeon, the response dependency on absolute stimulus size makes it unlikely.

Collision Detectors

Although encoding time to collision requires considerable response specificity, the requirements of a "collision cell" are less severe. These detectors signal the approach of an object on a collision course with the observer, but do not need to accurately represent the time to that event. Such units would be useful in low-level avoidance behavior. Recently, it has been proposed that Orthopteran DCMD neurons are sensitive to such approaching targets (Rind and Simmons, 1992; Simmons and Rind, 1992). Unlike cells tuned to expansion in MSTd, these units are not sensitive to the global motion pattern of the stimulus but detect more local features of the approaching object. Two cues present in looming stimuli are thought to be important in DCMD selectivity – the acceleration of image edges and the increase in edge length. These neurons are apparently not independently selective to divergence in the velocity field; no difference in response is observed between two fixed length edges whether they are moving towards or away from one another. Because DCMD units are specific for local stimulus attributes, positional invariance with regards to stimulus position within the receptive field is readily obtained. It is not believed that these cells encode tau because, like MST cells, their response is dependent on stimulus size as well as rate of expansion (Rind and Simmons, 1992).

DCMD cells respond well to expanding solid square stimuli but show less sensitivity for square outlines. In a study recently completed in our lab, we probed MST cells with similar patterns and obtained equally vigorous responses to each of these stimulus classes (see Chapter Two of this thesis). To account for their results, DCMD workers postulated the presence of separate channels to detect light/dark and dark/light transitions. The activities present in the two channels are at some point subtracted from one another. This explains the poor response of these cells to square outlines. The lines making up the edges of these stimuli consist of spatially adjacent dark/light and light/dark transitions. Real world stimuli more generally approximate the "solid square" case where these transitions within an object are relatively spatially separated. The purpose of this subtraction is so that randomly textured stimuli, composed of approximately equal dark/light and light/dark transitions, give a negligible response. In this way, specificity is obtained for detecting approaching objects as opposed to registering the approach of a surface because of the insect's own flight trajectory. MST cells lack this sort of specificity, and may well play a more generic role in representing motion pattern, as discussed next. In looking for "collision cells" in the primate, it may be worthwhile to consider subcortical areas of the brain serving as the sensory processors of this behavior, which is likely phylogenetically old.

Motion Pattern

Psychophysical studies have shown that correctly representing object motion through space, independent of self-motion, depends on detecting the size of the image of the target, the expansion or contraction of the object, and differences between the motion of the images on the two retinas (Brenner et al, 1993), all of which MST is sensitive to. It is easy to see that any rigid object that turns and/or rotates in depth will potentially introduce expansions, rotations, spirals, and translations into the flow. Consistent with these considerations, sporadically over the past decade investigators have raised the possibility that MST is involved in the analysis of the motion patterns objects generate as they move through space (Duffy et al, 1991). Graziano, et al (1994) developed this idea further, proposing a "motion pattern" hypothesis for the function of MSTd. They argued that these units are generically tuned for complex motion patterns in the environment and potentially can represent such stimuli as a rotating windmill, the expansion produced by a drop of water in a pond, and the complex motion generated by a monkey rotating an object in his hands. Selectivity to such varied stimuli is also a useful way of segmenting the environment into separate objects and surfaces moving in complex ways. This ideal has good theoretical appeal. Although area MT has been shown to be selective for local unidirectional (linear) motion, it is not clear how these local signals could be integrated into a cohesive perception of say, a cheetah running across a forest. It is possible that this integration is somehow occurring within MT itself, although how this would occur is not at all clear. Given that the heaviest forward connection of MT is to MST, this latter area would be an obvious choice for a location to "pool" local unidirectional motion cues into a unified percept. More complex motion patterns, such as a bird in flight, could be represented in MST by combining activities across several units tuned to different components of this motion. This "motion pattern" hypothesis does not preclude MSTd from playing a role in analyzing optical flow for the purpose of representing ego-motion. The retinal patterns produced as a consequence of observer motion can potentially be included within a very generic conception of motion pattern. In the final section of this review, we will discuss the plausibility of object motion and self-motion being processed within the same cortical system.

It has been standard practice to consider the response selectivity of units in MSTd in light of the computational and psychophysical literatures on ego-motion representation. The pattern motion literature, although not as rich, may also offer important insights into the physiology of this region. There are at least three possible general architectural frameworks with regards to a network that processes “motion pattern” information. The first is that such a system does not exist separately from a network of local motion detectors. The “pooling” referred to above would be an emergent property of a network of local motion detectors, such as those present in MT, and no explicit integration of signals need occur. Put simply, under this scheme the perception of motion pattern is no more than the perception of its component parts.

The second possibility is that the systems for processing local and pattern motion are completely independent. Characteristics of the two networks, such as detection thresholds and dynamic properties, would most likely be unrelated, except by chance. In other words, the characteristics of one system could not be predicted based on knowledge of the other. The final possibility is that the two networks are along different stages of the same pathway. As local information would be lost at the integration stage, such an arrangement would place the local unidirectional motion detectors prior to the motion pattern center. This is exactly what the anatomy and physiology for MT/MST suggests, and we consider this last alternative the most likely.

It may be argued that, based on single unit recording studies which found evidence for such integrative motion detectors in MSTd, that the issue of whether the nervous system contains a “motion pattern” system is a mute point. However, finding cells in the brain that respond to a particular stimulus does not mean that the organism necessarily has access to this information. In addition, the unit may be responding to an unexpected aspect of the stimulus. For example, there exist directionally selective cells in area V1 which are driven quite vigorously (and in a directionally selective way) by the complex flow patterns of expansion, rotation, and contraction. An expansion stimulus may drive such a cell quite well, while a contraction stimulus placed in the same spatial location gives no response. The experimenter might conclude that he had found cells in V1 tuned specifically to expansion. However, based on their equally vigorous responses to small patches of linear motion, it is not believed that individual V1 cells are directly representing these global motion patterns. The apparent selectivity for expansion was a consequence of the unit detecting motion over a small patch of this stimulus. Similarly, if the role of MSTd is limited to ego-motion perception, its units might still respond to the types of motion objects create while rotating and translating in depth, while not being directly involved with the representation of such information. The approach of our lab (see Chapter Two of this thesis) has been to define these complex motion patterns in a number of different ways. We believe that the form/cue invariance in preferred tuning we observed in MSTd argues in favor of a general role for this area in presenting complex motion patterns (CMPs). We will review psychophysical studies which are relevant to determining the level of interdependence between local unidirectional and pattern motion processing and attempt to reconcile this data with the anatomy and physiology. Pattern motion is sometimes referred to as "relative motion", stressing the fact that the perception of CMPs depends on detecting differences in component local motion orientation. We will argue that the two concepts are not identical and making this distinction is important for relating the psychophysical literature with the physiology of MST.

There is some perceptual evidence from human studies that a relative motion system does not exist. Two groups (Braddick and Holliday, 1991; Werkhoven and Koenderink, 1991) observed that the motion patterns of expansion, contraction, rotation, and shear do not "pop out" in displays that contain distracters. In these experiments, small patches of motion pattern are located randomly across a display. All these individual stimuli are identical, except for one stimulus which contains pattern motion of a different type. For example, in a display containing several regions of expansion pattern motion, one example of rotation is embedded. The task of the subject is to find this exception in a reaction time paradigm. As additional distracters are added, this task becomes increasingly more difficult. However, the relation between number of distracters and task completion time is an indication of how the information contained in the stimuli is processed. If the visual system can analyze this information from different spatial locations in parallel, search time increases much more slowly than if a “serial search” of each target is required. Stimulus attributes such as disparity, color, and spatial orientation have been shown to be processed in parallel as the exception will “pop out” from the field of distracters. Significantly, linear (unidirectional) stimuli also possess this property. A small patch of horizontally moving dots will be distinguished amongst a field of vertical distracters with a search time reflecting parallel processing of this information. However, as already mentioned, relative motion patterns do not possess this property. Search time for a rotating pattern in the presence of expanding distracters indicates that the human nervous system has to individually examine each pattern sequentially. This is true of all possible exception/distracter combinations of relative motion (expansion, contraction, two directions of rotation, shear, deformation). The authors of these studies argue that the human visual system lacks a network that specifically can process information about motion pattern. They argue that a selective attention mechanism is used to integrate these signals across local, unidirectional detectors (such as those found in MT).

To help clarify this line of reasoning, consider an example from engineering: A group of scientists have placed temperature sensors at different depths in the ocean. A readout of this information is available, giving instantaneous information about each of these different measurements. The temperature sensors process this information in parallel, because of separate systems responsible for each measurement. However, if a scientist needs information about temperature gradients across the different depths, he has to take the individual measurements and perform the required calculations by hand. If this scientist had arrays of such temperature detectors all across the ocean, he would have to calculate each of their gradients by hand, one at a time. If instantaneous measurements of these gradients were simultaneously required for all the arrays, an alternative system would have to be found. To accomplish this, each set of temperature measurements from an array could feed into separate computers that would do these computations, and display each gradient on a readout. Now the scientist could have immediate access to both individual temperatures and gradients from different parts of the ocean. Both these measurements would, in essence "pop out."

The analogy with linear detectors/pattern detectors and MT/MST should be clear. The fact that pattern motion does not "pop out" would argue against a specialized neural system, such as MST, playing a role in pattern motion perception. Instead, some attentional mechanism could be postulated to somehow pool this information from its distributed form in MT, just as the scientist in the scenario above had to calculate temperature gradients by hand. However, there is a potential weakness in this analysis. It needs be remembered that MT, but not MST, has a clear retinotopic organization. In the former region, arrays of linear motion detectors separately analyze local areas of the optical array. The large receptive fields in MST (which at times can cover the majority of the visual field) would prevent simultaneous computations of motion pattern in different parts of the visual field. Keeping this in mind, it would be reasonable to propose that while this region is involved is processing relative motion information, it can only handle one pattern at a time. Referring back to our analogy with the oceanographer, if he could afford only one computer rather than several to compute gradients, he would have to process the information across temperature arrays in a serial manner, despite having a system specifically dedicated to this task. If this is a valid analogy of how MST is processing motion pattern information, we would expect this area to show strong attentional effects. An interesting experiment would be to record from this region, while the monkey performed the "pop out" task for motion pattern outlined above. We would expect that a cell selectively tuned for the pattern of the exception stimulus to start firing as the monkey located this target.

Another possible objection to the proposal that the primate visual system contains a region specifically involved in motion pattern processing comes from psychophysical experiments on the perception of motion behind apertures. If a square rotates behind an aperture, such that the square's corners are not visible throughout the rotation, no coherent perception of rotation is achieved, and the stimulus appears to consist of non-rigid motion. Similarly, if a line is rotated around an axis perpendicular to the observer's line of sight, subjects are poor at locating its center of rotation when the ends of the line are not visible (Shiffrar, 1991). The center of rotation always appears to be located along the line, regardless of its actual location. Subjects do not appear to use global or high-level rigidity constraints in their interpretation of the global rotational motion present in these stimuli. It is important to recognize that the veridical interpretation of these stimuli is unambiguously specified if spatial pooling of information in accordance with some rigidity constraint is imposed. Intuitively, one might expect that "motion pattern" units with large receptive fields and considerable spatial integration, such as those proposed for MST, would be constructed to veridically represent rotation behind apertures. Since, instead, the perception of rotation seems constrained by the local information available to local motion detectors such as those present in MT, it seems unnecessary to postulate another cortical area to represent this motion pattern.

The fact that some cells in MT can successfully integrate information from differently oriented features that fall within their receptive fields (Stoner et al, 1992) but that additional integration fails to occur for more complex patterns is interesting, but still leaves open a role for an area such as MST to pool this information together, without imposing any additional satisfaction of constraints. The phenomena discussed above with regards to rotation behind apertures merely put limits on the spatial integration capabilities of a motion pattern system, rather than ruling out its existence. These psychophysical findings do suggest that the response selectivity of cells in MST may well be largely determined by the pattern of connectivity between MT and MST, rather than arising de nova from local interactions within MST itself. The failure to properly pool together the local signals present in these stimuli into a single rigid percept apparently reflects a failure to unambiguously represent the local motion of the line segments, do to the aperture problem in MT. The veridical representation of the motion of the corners of the square (or the ends of line segments), when visible, does not suffer from this problem because motion signals are available from nearby features containing at least two orientations. At least in theory, assuming rigid motion, the unambiguous linear motion of an object can be recovered by measuring the local motion of two differently oriented features and, through a process of constraint satisfaction, recover the motion of the object uniquely consistent with both these measurements (Hildreth et al, 1987). It should be pointed out that some additional integration of motion information may be occurring in MST, but that lines and squares rotating behind apertures provides a particularly challenging problem. It would be interesting to attempt to correlate the coherence/independence of these patterns with single unit activity in MST.

If the systems for detecting local unidirectional motion and pattern motion were completely independent, speed discrimination thresholds for these motion types would be the same only by chance. Psychophysical studies have shown that the thresholds for looming, rotation, and linear motion are all similar (Sekuler et al, 1992). How can these results be explained in light of our proposal that MST is specialized for analyzing pattern motion? It should be remembered that, as well as being selective for radial motion, the majority of units in MST are selective for linear translation as well. These cells are selective for the motion of an object moving across an observer's field of view. Although this is an example of unidirectional motion, we should still consider this a subtype of "pattern" motion because the perception is dependent on the integration of local motion signals, in this case all oriented in the same direction. This is consistent with what is known about the population of cells in MSTd. While there are cells tuned to radial type patterns, more than half of the cells in this region respond to the unidirectional motion of large patterns.

Rather than thinking of MST as a relative motion center, its role in pooling together local signals should be emphasized, which is why we refer to it as a “pattern” motion integrator. Psychophysical experiments which attempt to infer the existence of a relative motion system comparing linear and radial stimuli miss this distinction. The result that speed discrimination thresholds for translation, rotation, and expansion are similar does argue against distinct channels for processing unidirectional and relative motion. However it does not argue against a system specialized for pattern motion integration. This is completely consistent with single unit recording studies in MST, which have found both radial and linear motion tuning, even within the same unit. One study showed that superimposing (transparently) a linear motion field over an expanding random dot pattern generates an illusory shift in the perceived focus of expansion in the direction of the linear motion (Duffy et al, 1993). Although this argues that linear and radial motion channels are not independent, it does not argue against a “pattern” motion center.

Sekuler, et al (1992) argue that detection sensitivities for motion patterns can be predicted from sensitivities to the component local linear motions that make up the stimuli. However, different predictions should be made, depending on whether or not a separate system is available for analyzing relative motion. If such a pattern motion system does not exist, they argue for a simple pooling rule: Perceptual performance should increase with the square root of the number of independently processed motion elements. If there is a separate channel for such pattern motion processing, a complex pooling rule should be used to predict relative motion speed sensitivities from unidirectional thresholds. With complex pooling, relative motion sensitivities increase linearly with the number of unidirectional motion signals responsible for the global percept. Since their perceptual data demonstrated that the simple pooling scheme is more consistent with observed performance, the investigators favored a "common channel" model. Implicit in the design of these studies was the local vs. pattern distinction we have tried to emphasize. However, it should be pointed out that these pooling predictions are based on an ideal pattern motion system and that actual performance is likely to fall short of this. Rather than arguing against pattern motion sensors per se, we believe these findings instead put limitations on the performance of these detectors. As emphasized in the previous discussion of the perception of rotation behind apertures, if the response characteristics of units in MST are largely determined by the pattern of connectivity with MT rather than arising de nova in MST through local interactions, it is not surprising that their information content is limited by the local motion signals encoded in this earlier representation. We believe that both of these studies severely constrain possible mechanisms for generating response selectivity in MST.

The story becomes even more complicated when considering the results of perceptual fatigue studies (Regan, 1986; Regan and Beverley, 1978). Investigations that have compared relative and unidirectional motion channels using this approach have concluded that there are separate channels for expansion, contraction, rotation, and linear motion. The basic paradigm is as follows: Human subjects are exposed for an extended period of time with one type of motion stimulus, e.g. expansion. After this conditioning they are immediately asked to detect some component of motion in a test stimulus. The amount of motion in this stimulus is varied so that assessment of a detection threshold can be made. This threshold can then be compared with thresholds for this motion type in the absence of a conditioning period. The results show that different pattern motion channels can be fatigued separately. An expanding stimulus elevates motion detection thresholds for expansion but not for rotation, linear motion, or contraction. A similar independence exists for each combination of motion types, suggesting an independence of the channels representing these different stimuli. Based on somewhat different considerations using a masking paradigm, another group (Freeman and Harris, 1992) has also argued for the presence of separate relative motion processing systems.

The findings above are difficult to reconcile with our previous discussion of the coexistence of linear and radial specificity within the same MST units. To make things worse, double and triple component units have been characterized in our lab, as well as in others, demonstrating considerable mixing of motion pattern signals within the same units. A possible explanation is that partial segregation of these channels still exists in MST, as indicated by the presence of single component cells that are selective for a single stimulus type. Perhaps the results of these fatigue experiments can be explained by the contribution of this subclass of units on perception. We should mention that although conflicting with published data from our lab (Graziano et al, 1994), Lagae, et al (1994) reported finding no double-component cells that responded to both global divergence (expansion/contraction) and curl (rotation), suggesting an independence of these channels.

To sum up, we believe the existing evidence from psychophysics is consistent with a role for MST in pooling together local motion signals into coherent percepts. As such, it is an appropriate candidate for a region analyzing, among other stimuli, the motion of objects in the environment. The representation of these motion patterns, in terms of imposing constraint satisfactions such as rigidity, is largely determined by the local motion representation present in MT. In MST some segmentation of motion pattern into separate channels is preserved, but there is considerable interaction between these information streams.

Some single unit recording studies (Tanaka et al, 1989ab) obtained poor responses in area MST to stimuli smaller than 20 degrees. This would seem to make MST an unlikely candidate for the general analysis of motion pattern within the environment, as objects rarely subtend this large a visual angle. These results are likely a consequence of the anaesthetized conditions under which these experiment were conducted. Working under awake, behaving conditions, we have recorded brisk responses using stimuli as small as 5 degrees in diameter (unpublished observation). Stimuli even smaller than this may effectively drive MSTd units if the monkey is trained to attend to these patterns, such as in a discrimination task. Although not studied systematically, the presence of "rotation in depth" (rotation about the axis orthogonal to the line of sight) units reported by some workers offers further support for a motion pattern theory.

In a recently completed study by our group (see Chapter Two of this thesis), we tested the motion pattern hypothesis by generating CMPs using a number of different feature types. Besides the traditional random dot patterns employed in previous MST investigations, we recorded from the awake, behaving monkey using stimuli that more resembled objects moving in the environment, e.g. rotating and expanding squares. We also used cues other than luminance borders to produce "second-order" or non-fourier motion. Previous studies had given only a cursory consideration to feature invariance with regards to MST. These earlier studies looked at the effect of luminance boundary polarity (white stimulus on black background vs. black stimulus on white background) as well as varying the spatial texture of the stimuli. Consistent with our results, a relatively high degree of insensitivity to these parameters was reported.

Our experiments indicate that the preferred motion pattern for a majority of cells in area MST is not dependent on the features that define the motion. When statistically significant differences in preferred tuning existed, the magnitude of these differences tended to be small, compared with the range of preferred tunings present in the population. In general, the more robust the neuron's response to motion pattern, the smaller these differences were. Other parameters of the tuning curves, e.g. response height, did vary depending on the features used to define the motion pattern of the stimulus. Apparently, in terms of stimulus selectivity, MST extracts information related to the motion pattern of the stimulus, ignoring other stimulus attributes. We have proposed an analogy between MSTd and the inferotemporal region (area IT) to describe this finding (Graziano, et al 1994; see Chapter Two of this thesis).

Area IT has long been thought to analyze the spatial organization of stimulus features. In this region, investigators have reported the existence of "face" cells "toilet brush cells" and the like. Form/cue invariance has been reported in area IT as well (Sary et al, 1993). These investigators recognized that the perception of shape is invariant with respect to location in space, size, and the cues which define the shape. Using cues based on differences in luminance, motion, and texture, they found in IT a physiological correlate of this perceptual invariance. The neurons ignored aspects of the stimulus unrelated to spatial structure. We have discussed evidence that a analogous specificity for motion pattern has been delegated to area MST. The outputs of these two regions converge in parietal cortex, perhaps to pool together information from these different processing streams.

The form/cue invariance present in areas IT and MST suggests that the functions these two cortical regions play are potentially quite broad. Furthermore, the information represented in the motion pattern (MST) and spatial pattern (IT) pathways is, to a large extent, orthogonal. The significance of form/cue invariance in visual information processing was well expressed in Stoner et al (1992) with reference to the representation of linear motion: "The direction and rate at which an object moves are not normally correlated with the manifold physical cues (for example, brightness, and texture) that enable it to be seen. As befits its goals, human perception of visual motion largely evades this diversity of cues for image form; direction and rate of motion are perceived in a fashion that does not depend on the physical characteristics of the object." The reason this quality has emerged repeatedly in the brain is because of the computational efficiency is affords. As an analogy, consider our system of mathematics. We have one system to manipulate any kind of quantity, whether it be number of dogs, birds, or golf balls. The numerical computations we perform on these things do not depend on what the numbers are representing. Imagine having to use a different set of mathematical rules for every type of thing in the world. By breaking any analysis down into orthogonally separate parts, we can avoid elements of redundancy. If the brain had a separate system for the motion analysis of squares, circles, triangles, and what not, the brain would be prohibitively large. Thus, motion analysis occurs independently of the cues/features used to define this motion. Looked at from the opposite point of view, knowing that the brain exploits this strategy, we can design experiments that attempt to discover which axis of the stimulus "attribute space" the region is sensitive to.

By finding axes along which no differential response is detected, we further narrow our understanding of which stimulus characteristics are being represented. This notion of attribute orthogonality is commonly referred to in terms of "labeled lines." In this conception, perceptual systems are thought of as being broken down into separate channels. Each channel can be thought of as a separate axis in a multidimensional attribute space. This space is defined prior to the introduction of any particular stimulus. Every possible stimulus that we can perceive falls somewhere in this space. It is the combination of activity in each of these channels which determines where in the space the stimulus lies and what is perceived. By having these axes as independent (orthogonal) as possible, we not only reduce redundancy, but prevent interference between different processing domains which need to remain distinct. The orthogonality of MST and IT selectivity is an excellent example of exploiting this strategy.

These concepts have received attention in the psychophysical literature. One investigation (Stoner et al, 1992) determined that coherency rules with respect to plaid stimuli depend on component contrast and spatial frequency content but are independent of the cues that define these components. The rules underlying plaid coherency are thought to be the same as those used by the nervous system to parse local motion signals up into objects. Another study (Berkeley et al, 1994) demonstrated that the well known perceptual illusion of the tilt aftereffect persisted despite the conditioning and test stimuli having their oriented contours defined by different cues. Interestingly, the relative saliency of these cues was found to be important, but cue invariance was reported as long as saliency was controlled for. Single unit recording studies in area V1 of the macaque have demonstrated that these neurons are form/cue invariant in their selectivity for image contours (Albright et al, 1989) and in MT for local linear motion signals (Albright, 1992). This independence held for bandwidth and preferred tuning, but not for response amplitude. The average difference in preferred tuning direction using different stimulus cues did not vary from zero, and no more than 30 percent of the differences were outside the range of +- 45 degrees. This is similar to our findings in MST for motion pattern (see Chapter Two of this thesis).

We have suggested that MST is representing motion pattern and is effective at extracting this attribute from diverse groups of stimuli. We have discussed this selectivity in terms of the concept of "labeled lines" and "form/cue" invariance. The fact that objects moving in the environment can produce similar patterns of motion on the retina as are produced by observer translation has been discussed. We believe that a consideration of the available evidence leaves it an open question whether or not MST is representing motion from these two different sources. We will now look at this issue from another point of view: What is the evidence that these two types of motion are processed together in brain? If the evidence is strongly against coprocessing, we can further develop models and design experiments in the context of distinguishing between two mutually incompatible functions.

Ego-motion vs. Object Motion

In the discussion above of the significance of the back projections from MST to MT, considerable attention was paid to the importance of distinguishing motion in the environment from motion as a consequence of observer translation/rotation relative to the environment. Other authors, both in the psychophysical and computational literatures, have emphasized the importance of being able to separately extract this information for survival (Swanston et al, 1987; Hildreth et al, 1992). Hildreth (1992) in a paper which refines Rieger's differential motion model for direction of heading determination, discusses the importance of discounting motion in the environment before doing DOH computations. She stresses discontinuities in the velocity field as being important in object segmentation, an important first step in determining which environmental motion to discount. A different approach to this same issue is a series of models under the heading Focal/Ambient theory. Advocates of this distinction argue that there are two modes of processing visual information. "Focal" processing is involved in the analysis of local features in the environment, such as the shape or motion of an object. "Ambient" information processing involves the analysis of the visual scene as a whole, ignoring local features of this environment. The type of calculations thought to be used in ego-motion representation would fall under this latter category.

The focal/ambient distinction has some theoretical and intuitive appeal. For example, while analyzing optical flow data, local disruptions of the flow field, as caused by object motion, need to be ignored before DOH computations are made. During object motion analysis, these local disruptions need to be isolated from the rest of the flow data before this local information can be processed. In terms of our own exploring of the visual scene, there is a clear trade-off between processing focal and ambient information. When carefully studying part of the world, say when reading a book, vigilance with regards to the rest of the visual scene is compromised. In addition, there is psychophysical evidence that these two types of behaviors may use different anatomical regions or the same region in a different mode (Andersen, 1986). If such a distinction is valid with regards to object motion vs. global optic flow analysis, it may be difficult to place both of these functions in area MST. However, this would not necessarily rule out both object motion and ego-motion having MST in common. It is quite possible that MST is a common point in both processing streams and that later these two pathways diverge. The Focal/Ambient distinction merely requires a separation of function at some stage in processing. This need not even be a spatial segregation; it is possible that a single brain region is doing both types of processing, but at different times.

There is some evidence that object motion and optical flow are processed in common. Blurring the visual scene (equivalent to a low pass filter) effects response latencies for object motion, postural balance, and self motion equally, suggesting a shared common pathway (Straube et al, 1990). Similarly, patients with minimal cortical visual functions for object motion perception (Riddoch phenomenon) also have reduced visual self-motion perception. These studies do not necessarily argue for complete coprocessing of information from these two modalities, but do imply that there is some commonality in their streams of analysis. Another study by Kruk and Regan (1983) found that the ability of pilots to land a plane was correlated with speed discrimination for looming objects. Interestingly, thresholds for object motion have been found to be raised during periods of perceived self-motion, whether self-motion was induced by visual, vestibular, or somatosensory stimulation (Probst et al, 1986; Wertheim, 1981). These last two studies argue for a fairly close association with regards to processing ego-motion and object motion.

Much evidence also exists for a certain degree of independence between the processing streams for ego-motion and object motion. One human perceptual study (Brenner, 1991) concluded that the speed of the background, independent of whether the observer thought that he was moving relative to the scene as a whole, affected the perceived speed of an object moving over this background. Apparently this correction for background motion is a local phenomenon and does not depend on a global representation of observer motion, such as that obtained from ego-motion processing. Nakayama (1985) in a review of human visual information processing discusses the possibility of having two separate motion-analysis systems. In one conception of this idea, the AOS is identified as analyzing averaged motion over the whole visual field, while a cortical system is used to examine local regions of the environment. If this hypothesis is to believed, area MST would be a likely candidate for object motion perception and the AOS for self-motion perception. The fact that MST sends projections to the AOS may explain some of the psychophysical interactions between object and self-motion perception. In addition, we have previously discussed that some exchange of information is required between these two systems, if only to indicate which parts of the velocity field should be disregarded for their respective functions.

These ideas have a precedence in research on lower animals. A number of studies investigating the behavior and physiology of the fly's visual system indicate that this animal has two separate networks for analyzing the motion of objects and self-motion (Egelhaaf, 1988; Egelhaaf and Borst, 1993). The two systems differ in their interocular interactions, dynamic properties, as well as their spatial and temporal integration characteristics. Because of their different response properties, these two systems are not thought to interact much under natural conditions. Some authors have proposed that the fly is capable of detecting three distinct types of motion present within the retinal flow (Egelhaaf and Borst, 1993): large-field motion such as that produced from course deviation; image expansion occurring when the animal approaches an obstacle; and the relative motion of an object against a background. Separate neural networks specializing in the analysis of each of these motion types have been postulated to control the respective behaviors of course stabilization, landing behavior, and the fixation of objects. Note that in our consideration of the primate visual system, we have lumped together the first two perceptual/motor systems into a single ego-motion network. Recording studies on the fly have identified separate neurons selective for these different motion types. Global motion signals are thought to be represented in the horizontal cells present in the lobula plate while figure cells are believed to be sensitive to the motion of small objects.

The bulk of evidence from modeling, psychophysics, and studies of lower animals seems to be in favor of at least a partial separation for the processing of object and self motion. The exact nature of this independence is not clear. In the extreme case, there would be completely independent systems for the processing of these information types, perhaps diverging as early as the retina into geniculo-striate and AOS streams. Alternatively, the two pathways could diverge much later, with the final representations of ego-motion and object motion occurring in distinct cortical regions. Allowing for even greater coprocessing, a single cortical region such as MST could be thought of as a generic "pattern motion" detector with the self versus object motion distinction never being made explicit, but rather present in the distributed representation of this area. Several other roles for MST were also discussed including vection, structure from motion, and "reference signal." The available evidence cannot rule out a role for this region in processing such "optical flow" related information, although evidence was discussed (for example, with regards to vection) arguing against MST being the final stage in this process. We believe it is likely that MST is a least part of the processing stream involved in recovering data encoded in the positional velocity field of optical flow. As such, it is prudent at this juncture to characterize MST as a generic "pattern motion" integration center which can send this information to other regions for further processing, as required. Unfortunately, the cues which help to distinguish between the various motion sources have not been well worked out. Or rather, they have been well worked by our perceptual system, but our cognitive systems have yet to figure out how this process occurs. Because these cues may be fairly subtle, physiologist should work on making their stimuli more ecological when studying phenomena such as ego-motion and object motion.

REFERENCES

Albright TD (1984) Direction and Orientation Selectivity of Neurons in Visual Area MT of the Macaque. Journal of Neurophysiology 52(6): 1106-1130.

Albright TD (1989) Centrifugal direction bias in the middle temporal visual area (MT) of the macaqu. Visual Neuroscience 2:177-188.

Albright TD (1992) Form-Cue Invariant Motion Processing in Primate Visual Cortex. Science 28: 1141-1143

Albright TD, Chaudhuri A (1989) Soc. Neurosci. Abstr. 15: 323.

Andersen GJ (1986) Perception of Self-Motion: Psychophysical and Computational Approaches. Psychological Bulletin 99(1): 52-65.

Andersen GJ, Braunstein ML (1985) Induced self-motion in central vision. Journal of Experimental Psychology: Human Perception and Performance 11:122-132.

Andersen RA (1989) Visual and eye movement functions of the posterior parietal cortex. Annu Rev Neurosci 12:337-403.

Ansorge K, Grusser-Cornehls U (1977) Visual and visual-vestibular responses of frog cerebellar neurons. Exp. Brain Res. 29:445-465.

Berkley MA, Debruyn B, Orban G (1994) Illusory, motion, and luminance-defined contours interact in the human visual system. Vision Res 34(2): 209-16.

Beusmans JM (1993) Computing the direction of heading from affine image flow. Biol. Cybern. 70(2):123-36.

Boussaoud D, Ungerleider LG, Desimone R (1990) Pathways for motion analysis: Cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. Journal of Comparative Neurology 296:462-496.

Braddick OJ, Holliday IE (1991) Serial search for target defined by divergence or deformation of optic flow. Perception 31:345-354.

Brenner E (1993) Judging an object's velocity when its distance changes due to ego-motion. Vision Res 33(4): 487-504.

Brenner E Judging Object Motion During Smooth Pursuit Eye Movements: The Role of Optic Flow. Vision Research 31(11): 1893-1902.

Bruce C, Desimone R, Gross CG (1981) Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. J. Neurophysiology 46:369-385.

Cavallo V, Laurent M (1988) Visual information and skill level in time-to-collision estimation. Perception 17(5):623-32.

Desimone R, Albright TD, Gross CG, Bruce CJ (1984) Stimulus selective properties of inferior temporal neurons in the macaque. J Neurosci 4:2051-2062.

Desimone R, Ungerleider LG (1986) Multiple Visual Areas in the Caudal Superior Temporal Sulcus of the Macaque. The Journal of Comparative Neurology 248: 164-189.

Duffy CJ, Wurtz RH (1991a) Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. Journal of Neurophysiology 65(6): 1329-1345.

Duffy CJ, Wurtz RH (1991b) Sensitivity of MST neurons to optic flow stimuli. II. Mechanisms of response selectivity revealed by small-field stimuli. Journal of Neurophysiology 65(6): 1346-1359.

Duffy CJ, Wurtz RH (1993) An illusory transformation of optic flow fields. Vision Res 33(11): 1481-90.

Egelhaaf M, Hausen K, Reichardt W, Wehrhahn C (1988) Visual couse control in flies relies on neuronal computation of object and background motion. Trends in Neuroscience 11(8): 351-358.

Egelhaaf M, Borst A (1993) Motion computation and visual orientation in flies. Comp. Biochem. Physiol A. 104(4):659-73.

Erickson RG, Thier P (1991) A neuronal correlate of spatial stability during periods of self-induced visual motion. Experimental Brain Research 86: 608-616.

Fluckiger M, Baumberger B (1988) The perception of an optical flow projected on the ground surface. Perception 17(5):633-45.

Freeman TC, Harris HG (1992) Human sensitivity to expanding and rotating motion: Effects of complementary masking and directional structure. Vision Research 32: 81-87.

Frost BJ, Wylie DR, Wang YC (1990) The processing of object and self-motion in the tectofugal and accessory optic pathways of birds. Vision Resesearch 30(11):1677-1688.

Ghelarducci B, Ito M, Yagi N (1975) Impulse discharges from flocculus Purkinje cells of alert rabbits during visual stimulation combined with horizontal head rotation. Brain Res. 87:66-72.

Gibson JJ (1950) The perception of the visual world. Boston: Houghton Mifflin.

Gielen CC, van Asten WN (1990) Postural responses to simulated moving environments are not invariant for the direction of gaze. Exp Brain Res 79(1):167-74.

Graziano MS, Andersen RA, Snowden RJ (1994) Tuning of MST neurons to spiral motions. Journal of Neuroscience, 14, 54-67.

Gross CG, Rocha-Miranda CE, Bender DB (1972) Visual properties of neurons in inferotempral cortex of the macaque. J Physiol (Lond) 35:96-111.

Hietanen JK, Perrett DI (1993) Motion Sensitive cells in the macaque superior temporal polysensory area. I. Lack of response to the sight of the animal's own limb movement. Exp Brain Res. 93(1): 117-28.

Hikosaka K, Iwai E, Saito HA, Tanaka K (1988) Polysensory properties of neurons in the anterior bank of the caudal superior temporal sulcus of the macaque monkey. J. Neurophysiol. 60:1615-1637.

Hildreth EC (1991) Recovering heading for visually guided navigation. Vision Research 32(6): 1177-1192

Hildreth EC, Koch C (1987) The Analysis of Motion: From Computational Theory to Neuronal Mechanism. Annual Review of Neuroscience 10: 477-533.

Judge SJ (1990) Vision. Knowing wehre you're going [news; commend]. Nature 348(6297).

Kellman PJ, Gleitman H, Spelke ES (1987) Object and observer motion in the perception of objects by infants. J Exp Psychol Hum Percept Perform 13(4):586-93.

Koenderink JJ (1986) Optic flow. Vision Research 26(1):161-180.

Koenderink JJ, van Doorn AJ (1981) Exterospecific component of the motion parallax field. Journal Optical Society of America 71(8): 953-957.

Komatsu H, Wurtz RH (1988) Relation of cortical areas MT and MST to pursuit eye movements. III. Interaction with full-field visual stimulation. Journal of Neurophysiology 60(2): 621-644.

Komatsu H, Wurtz RH (1988) Modulation of pursuit eye movements by stimulation of cortical areas MT and MST. Journal of Neurophysiology 62(1):31-47.

Kruk R, Regan D (1983) Visual test results compared with flying performance in telemetry-tracked aircraft. Aviation, Space, and Environmental Medicine 54:906-911.

Lagae L, Maes H, Raiguel S, Xiao DK & Orban, GA (1994). Responses of macaque STS neurons to optic flow components: a comparison of areas MT and MST. Journal of Neurophysiology 71:1597-626.

Lappe MR, Rauschecker JP (1993) A neural network for the processing of optic flow from ego-motion in man and higher mammals. Neural Computation 5:374-391.

Larish JF, Flach JM (1990) Sources of Optical Information Useful for Perception of Speed of Rectilinear Self-Motion. Journal of Experimental Psychology: Human Perception and Performance 16(2): 295-302.

Lisberger SG, Morris EJ, and Tychsen, L (1987) Visual motion processing and sensory-motor integration for smooth-pursuit eye movements. Annu. Rev Neurosci 10:97-129.

Longuet-Higgins HC (1986) Visual motion ambiguity. Vision Research 26(1):181-3.

Maioli MG, Squartrito S, Domeniconi R (1989) Projections from visual cortical areas of the superior temporal sulcus to the lateral terminal nucleus of the accessory optic system in macque monkeys. Brain Research 498: 389-392.

Masud H, Treue S & Andersen RA (1989) Surface Interpolation in Three-Dimensional Structure-from-Motion Perception. Neural Computation, 1:324-333.

Maunsell J, van Essen DC (1983a) Fuctional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. Journal of Neurophysiology 49(5): 1127-1147.

Maunsell J, van Essen DC (1983b) Functional properities of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitiviy to binoncular disparity. Journal of Neurophysiology 49(5): 1148-1167.

Mestre DR, Brouchon M, Ceccaldi M, Poncet M (1992) Perception of optical flow in cortical blindness: a case report. Neuropsychologia 30(9):783-95.

Motter BC, Mountcastle VB (1981) The functional properties of the light-sensitive neurons of the posterior parietal cortex studied in awake monkeys: foveal sparing and opponent vector organization. J. Neurosci. 1:3-26.

Nakayama K (1985) Biological Image Motin Processing: A Review. Vision Research 25(5): 625-660.

Nawrot M, Blake R (1993) Visual alchemy: stereoscopic adaptation produces kinetic depth from random noise. Perception 22(6):635-42.

Newsome WT, Wurtz RH, Komatsu H (1988) Relation of Cortical areas MT and MST to pursuit eye movements. II. Differentiation of retinal from extraretinal inputs. Journal of Neurophysiology 60(2): 604-620.

Orban GA, Lagae L, Verri A, Raiguel S, Xiao D, Maes H, Torre V (1992) First-Order analysis of optical flow in monkey brain. Proceedings of the National Academy of Science 89: 2595-2599.

Pailhous J, Ferrandez AM, Fluckiger M, Baumberger B (1990) Unintentional modulations of human gait by optical flow. Behavioral Brain Research 38:275-281.

Poggio T, Verri A, Torre V (1991) Green Theorems and Qualitative Properties of the Optical Flow (1991) MIT AI Memo 1289.

Prazdny K (1980) Egomotion and relative depth map from optical flow. Biological Cybernetics 36: 87-102.

Probst T, Brandt T, Degner D (1986) Object-motion detection affected by concurrent self-motion perception: psychophysics of a new phenomenon. Behav Brain Res 22(1): 1-11.

Regan D, Beverley KI (1978) Looming detectors in the human visual pathway. Vision Research 18: 415-421.

Regan D, Beverley, KI (1982) How Do We Avoid Confounding the Direction We are Looking and the Direction We are Going. Science 215: 194-196.

Regan D (1986) Visual processing of four kinds of relative motion. Vision Research, 26, 127-145.

Regan D, Giaschi D, Sharpe JA, Hong XH (1992) Visual processing of motion-defined from: Selective failure in patients with parietotemporal lesions. Journal of Neuroscience 12(6):2198-2210.

Regan D, Hamstra SJ (1993) Dissociation of discrimination thresholds for time to contact and for rate of angular expansion. Vision Research 33(4):447-62.

Rieger JH, Lawton DT (1985) Processing differential image motion. Journal Optical Society of America 2(2): 354-359.

Rind FC, Simmons PJ (1992) Orthopteran DCMD neuron: a reevaluation of responses to moving objects. I. Selective responses to approaching objects. J Neurophysiol. 68(5):1654-66.

Rosenberg AF, Ariel M (1990) Visual-response properties of neurons in turtle optic neucleus in vitro. Journal of Neurophysiology 63(5):1033-45.

Roy JP, Wurtz RH (1990) The role of disparity-sensitive cortical neurons in singalling the direcition of self-motion. Nature 348:160-163.

Roy JP, Komatsu H, Wurtz R (1992) Disparity Sensitivity of Neurons in Monkey Extrastriate Area MST. Journal of Neuroscience 12(7): 2478-2492.

Royden CS, Banks MS, Crowell JA (1992) The perception of heading during eye movements. Nature 360(6404): 583-5.

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(1): 145-157.

Sakata H, Shibutani H, Kawano K, Harrington TL (1985) Neural mechanisms of space vision in the parietal association cortex of the monkey. Vision Res 25:453-463.

Sakata H, Shibutani H, Ito Y, Tsurugai K (1986) Parietal cortical neurons responding to rotatory movement of visual stimulus in space. Experimental Brain Reseach 61:658-663.

Sary G, Vogels R, Orban GA (1993) Cue-invariant shape selectivity of macaque inferior temporal neurons. Science 260(5110): 995-7.

Schwartz EL, Desimone R, Albright T, Gross C (1983) Shape recognition and inferior parietal neurons. Proc Natl Acad Sci USA 80: 5776-5778.

Sekuler AB (1992) Simple-pooling of unidirectional motion predicts speed discrimination for looming stimuli. Vision Res 32(12): 2277-88.

Shiffrar M, Pavel M (1991) Percepts of Rigid Motion Within and Across Apertures. Journal of Experimental Psychology 17(3): 749-761.

Simmons PJ, Rind FC (1992) Orthopteran DCMD neuron: a reevaluation of responses to moving objects. II. Critical cues for detecting approaching objects. J Neurophysiol 68(5): 1667-82.

Simpson JI, Leonard CS, Soodak RE (1988) The accessory optic system. Analyzer of self-motion. Ann N Y Acad Sci. 545:170-9.

Steinmetz MA, Motter BC, Duffy CJ, Mountcastle VB (1987) Funtional properties of parietal visual neurons: radial organization of directionalities within the visual field. Journal of Neuroscience 7(1):177-91.

Stoner GR, Albright TD (1992) Motion Coherency Rules are Form-cue Invariant. Vision Research 32(3): 465-475.

Straube A, Brandt T (1987) Importance of the visual and vestibular cortex for self-motion perception in man (circularvection). Hum Neurobiol. 6(3):211-8.

Straube A, Paulus W, Brandt T (1990) Influence of visual blur on object-motion detection, self-motion detection and postural balance. Behavioural Brain Research 40: 1-6.

Suzuki DA, May JG, Keller EL, Yee RD (1990) Visual motion response properties of neurons in dorsalateral pontine nucleus of alert monkey. Journal of Neurophysiology 63(1):37-69.

Swanston MT, Wade NJ, Day RH (1987) The representation of uniform motion in vision. Perception 16(2): 143-59.

Tanaka K, Hikosaka K, Saito H, Yukie M, Fukada Y, Iwai E (1986) Analysis of Local and Wide-Field Movements in theSuperior Temporal Visual Areas of the Macaque Monkey. Journal of Neuroscience. 6(1): 134-144.

Tanaka K, Fukada Y, Saito H (1989) Underlying Mechanisms of the Response Specificity of Expansion/Contraction and Rotation Cells in the Dorsal Part of the Medial Superior Temporal Area of the Macaque Monkey. Journal of Neurophysiology 42(3): 642-656.

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(3): 626-641.

Tanaka K, Sugita Y, Moriya M, Saito H (1993) Analysis of object motion in the ventral part of the medial superior temporal area of the macaque visual cortex. Journal of Neurophysiology 69(1):128-142.

Telford L, Spratley J, Frost BJ (1992) Linear vection in the central visual field facilitated by kinetic depth cues. Perception 21(3):337-49.

Turano K, Wang X (1994) Visual discrimination between a curved and straight path of self motion: effects of forward speed. Vision Res 34(1): 107-14.

Ungerleider LG Desimone R (1986a) Projections to the superior temporal sulcus from the central and peripheral field representations of V1 and V2. Journal of Comparative Neurology 248:147-163.

Ungerleider LG, Desimone R (1986b) Cortical Connections of Visual Area MT in the Macaque. Journal of Comparative Neurology 248: 190-222.

van Asten WN, Gielen CC, van der Gon JJ (1988) Postural adjustments by simulated motion of differently structured environments. Experimental Brain Research 73(2):371-83.

van den Berg, AV (1992) Robustness of perception of heading from optic flow. Vision Research 32(7): 1285-1296.

van den Berg AV (1993) Perception of heading [letter]. Nature 365(6446): 497-8.

van Essen DC, Maunsell JHR (1983) Hierarchical organization and functional streams in the visual cortex. Trends in Neuroscience 6: 370-375.

Waespe W, Henn V (1981) Visual-vestibular interaction in the flocculus of the alert monkey. II. Purkinje cell activity. Exp. Brain Res. 43:349-360.

Wang Y, Frost BJ (1992) Time to collision is signalled by neurons in the nucleus rotundus of pigeons. Nature 356(6366):236-8.

Warren WH, Blackwell AW, Kurtz KJ, Hatsopoulos NG, Kalish ML (1991) On the sufficiency of the velocity field for perception of heading. Biological Cybernetics 65: 311-320.

Warren WH, Hannon DJ (1988) Direction of self-motion is perceived from optical flow. Nature 10(336): 162-163.

Warren WH, Morris MW, Kalish M (1988) Perception of translational heading from optical flow. Journal of Experimental Psychology: Human Perception and Performance 14(4):646-660.

Warren WH, Hannon DJ (1990) Eye movements and optical flow. Journal Optical Society of America 7(1): 160-169.

Warren WH, Mestre DR, Blackwell AW, Morris MW (1991) Perception of circular heading from optical flow. Journal of Experimental Psychology: Human Perception and Performance 17(1): 28-43.

Werkhoven P, Koenderink JJ (1991) Visual processing of rotary motion. Perception and Psychophysics 49: 73-82.

Wertheim AH (1981) On the relativity of perceived motion. Acta Psychol. 48: 97-110.

Wetheim AH (1987) Retinal and extraretinal information in movement perception: how to invert the Filehne illusion. Perception 16(3):299-308.

Wurtz RH, Komatsu H (1988) Relation of cortical area MT and MST to pursuit eye movements. I. Localization and visual properties of neurons. Journal of Neurophysiology 60(2): 580-603.

Wylie DR, Frost BJ (1991) Purkinje cells in the vestibulocerbellum of the pigeon respond best to either translational or rotational wholefield visual motion. Experimental Brain Research 86(1): 229-232.

Wylie DR, Frost BJ (1993) Responses of pigeon vestibulocerebellar neurons to optokinetic stimulation. II. The 3-dimensional reference frame of rotation neurons in the flocculus. J Neurophysiol 70(6): 2647-59.

Wylie DR, Kripalani T, Frost BJ (1993) Responses of pigeon vestibulocerebellar neurons to optokinetic stimulation. I. Functional organization of neurons discriminating between translational and rotational visual flow. J Neurophysiol 70(6): 2632-46.

Yamasaki DS, Wurtz RH (1991) Recovery of function after lesions in the superior temporal sulcus in monkey. Journal of Neurophysiology 66(3):651-673.

Chapter 2

The Representation of Motion Pattern in Form/Cue Invariant MST Neurons of the Macaque

ABSTRACT

Several groups have proposed that area MST of the macaque monkey has a role in processing optical flow information relevant to the analysis of self motion. We propose that this cortical region may also be important in analyzing the complex motions that objects undergo as they move through the environment. Under this more general conception, MST is involved in the generic function of motion pattern representation, with its units responsible for "pooling" local motion signals sent forward from area MT into a more unified representation. A prediction of this hypothesis is that the preferred motion pattern for a particular neuron should remain invariant with regards to the features and cues which define this motion. Accordingly, we were interested in determining the diversity of stimulus classes over which these cells can extract information about complex motion patterns such as expansion, contraction, and rotation. To characterize the responses of MST cells, we recorded from 190 cells in the awake behaving monkey. We probed the receptive fields of these neurons with stimuli consisting of different classes of features and cues, but with the same panel of motion pattern types. Gaussian shaped tuning curves were constructed for each experiment based on the differential responses. The different classes of stimuli included coherently moving random dot patterns (190 cells tested), solid squares (184 cells tested), outlines of squares (119 cells tested), a square aperture moving in front of an underlying stationary pattern of random dots (119 cells tested), a square composed entirely of flicker (119 cells tested), and a square of non-fourier motion (50 cells tested). When a unit responded well to multiple stimulus classes, most cells ignored the spatial differences between stimuli, at least with regards to their preferred tuning. Although the magnitude of response often varied between classes, in those cells that responded strongly to multiple classes of stimuli, the orientations of the tuning curves were similar. The stronger the responses, the more similar the tuning curves. To perform our analysis, we introduce a novel index of response strength which employs techniques from statistical hypothesis

testing. Although preferred tuning direction remained remarkably invariant across stimulus classes, the amplitude and width of the response curves often varied considerably. In particular, the presence of flicker in the stimuli inhibited cell responsiveness. Also, the presence of stationary features within the patterns appeared to suppress a cell's response to the moving elements in the display. Despite these qualifications, MST is able to extract information about motion type for a wide range of stimuli, making it an appropriate candidate for analysis of environmental motion, as well as motion introduced by observer translation.

INTRODUCTION

Two pathways for visual information processing in extrastriate cortex have been identified (van Essen and Maunsell, 1983). One stream, the "what" pathway, which sends information ventrally into the temporal lobe, appears to be involved in processing the spatial pattern of the visual scene. The other stream, projecting dorsally into posterior parietal cortex, has been described as the "where" pathway and is involved in localizing objects in space and the related task of processing the motion present in an image. The best studied area with regards to this latter task is area MT, located on the posterior bank and floor of the superior temporal sulcus (STS), whose cells have been shown to respond to simple linear (translational) motion (Maunsell and van Essen, 1983ab; Albright, 1984). MT sends a heavy projection forward to area MST, an adjacent cortical region located on the floor and anterior bank of the STS. Area MST (medial superior temporal region), which is also believed to play an important role in the motion processing hierarchy, has been functionally segmented into at least two distinct regions, a ventral lateral one (MSTl) and a dorsal one (MSTd) (Desimone and Ungerlieder, 1986; Saito, et al, 1986; Ungerleider and Desimone, 1986ab; Komatsu and Wurtz, 1988). The cells in MSTl have been shown to have relatively small receptive fields (4 degrees), rather similar in size to those found in area MT at the same eccentricity and also similar in terms of their preferrence for translational motion. Cells in MSTd, on the other hand, have rather large receptive fields

(>40 degrees), are often bilateral in their expanse, and generally include the fovea. These cells are specific not only for large-field translational motion but for stimuli that have expanding, contracting, and rotating elements (Sakata et al, 1985; Saito et al, 1986; Tanaka et al, 1986, 1989; Tanaka and Saito, 1989). Patterns of this type are sometimes referred to as elementary flow components, or EFCs, in recognition of the fact that even more elaborate motion patterns can be considered as linear combinations of these basic flavors. Our lab's previous investigation in MST showed that some units in this region have even more complex response characteristics, in many cases demonstrating a preference for spiraling motion pattern over either expansion, contraction or rotation (Graziano et al, 1994).

Because these complex motion patterns are built up from local regions of approximately straight motion, large EFC stimuli can effectively drive many units in MT lacking inhibitory surrounds. What distinguishes a large proportion of MSTd cells, besides the increased size of their receptive fields, is that their specificity for motion pattern is largely invariant with regards to the placement of the stimulus within the receptive field (Duffy and Wurtz, 1991ab; Graziano et al, 1994). We refer to this property as positional invariance. This invariance is with respect to preferred tuning; this property is less pronounced with respect to response amplitude and tuning width. Tuning invariance with respect to EFCs is never found within MT, where even minor positional shifts in the placement of the these stimuli will dramatically alter (even reverse) a unit's preferred tuning (Lagae et al, 1994).

The types of motion pattern MST cells respond to have been associated with the full field patterns projected onto the during locomotion, as first characterized by Gibson (1950). This suggests a role for MST in processing ego-motion and determining direction-of-heading. Many computational and psychophysical studies have shown that, by analyzing these flow field motion patterns, the parameters of observer rotation and translation can be recovered (Andersen, 1986; Prazdny, 1980; Warren and Hannon, 1990). The well documented positional invariance of MST units is puzzling — we would expect that stimulus placement should have a dramatic effect on unit response selectivity if direction of heading were being represented in a straight forward way. For example, in the absence of head and eye rotation, the global focus of expansion in the optical flow field is located in the direction of heading (DOH). If the nervous system used optical flow to guide navigation, we would expect neurons performing DOH analysis to be sensitive to the placement the focus of expansion. It is also unlikely that positional invariance emerged by chance — the connectivity required for this property to emerge is most non-trivial and has yet to be worked out.

An alternative proposal for MST, introduced in a previous paper (Graziano et al, 1994) makes an analogy between MST and area IT in the temporal lobe. Where IT is thought to analyze spatial pattern information in the image, MST could analyze motion pattern information. Cells in IT have been found that are selective for such complex spatial patterns as toilet brushes and faces (Gross et al, 1972; Desimone et al, 1984). This selectivity is maintained regardless of stimulus placement within the unit's large receptive field (Schwartz et al, 1983; Desimone et al, 1984). The positional invariance in cell tuning for both IT and MST suggests a functional connection between the two areas. It should be pointed out that the possible "pattern" motion and "ego-motion" roles for MST are not mutually exclusive and that both types of information could potentially be processed in this one region.

Most experiments studying area MST have used random dot stimuli with different types of global motion (expansion, contraction, translation motion, etc.) to probe the response properties of these cells. In the current investigation we have extended this approach to include stimulus types whose motion pattern is established by features other than random dots, such as object edges, and then compare these responses and tuning curves across classes. We also explore the effect of using cues other than luminance edges to define these features through stimuli created using "second-order" or non-fourier motion. These experiments will help to establish how general are the feature-types MST uses to extract motion pattern. This work is partially motivated by studies of "form/feature/cue invariance" recently demonstrated in MT, V1, and IT (Albright, 1992; Albright and Chaudhuri, 1991). It is also relevant to the discussion of "labeled lines" of information processing in cortex.

METHODS

Animal preparation

Two Rhesus monkeys were used for these experiments. The first monkey we studied (89-1) provided data from a single hemisphere; the other hemisphere was used for an unrelated experiment. We collected data from both hemispheres of a second monkey (90-2). The first set of pilot experiments were done with 89-1 and initially involved optimizing parameters such as stimulus size, type, speed, etc. Only data comparing the tuning curves of solid squares and random dots was collected from this monkey. Many more experiments were done with 90-2 than with 89-1. Because the results were similar in the two monkeys, the data was pooled for the purpose of analysis. Units located in MSTd were tentatively identified based on their location in the chamber and depth relative to the dura. In each of the three chambers recorded from, we mapped out both MSTd and MT based on the tuning characteristics of cells in these regions. Particularly helpful in distinguishing of MSTd from MT was the former cells' large receptive fields and positional invariance with respect to stimulus location in the receptive field. If we were not sure that a unit was within MSTd, we excluded data recorded from this neuron from further study. Based on these initial criteria, 190 cells were considered for this analysis, 119 from 90-2 and 71 from 89-1. All recordings were carried out under awake, behaving conditions. A fixation task (see below) was used to maintain a stable retinal image of the stimulus. A scleral search coil and an acrylic skull cap were implanted into the monkey five days before training on the fixation task began. Training and subsequent behavior were reinforced by depriving the monkeys of fluid before each session and then giving drops of apple juice upon correct task completion. Following mastery of the behavior, a second surgery was performed to provide a craniotomy that allowed chronic access to the brain for recording purposes. Around the craniotomy, a plastic chamber was attached to allow fixation of a Narashige microdrive which permitted fine control of electrode position. For these experiments we used varnish coated tungsten electrodes in conjunction with a guide tube to protect the electrode as it passed through the dura. All three chambers attached to the two monkeys were placed vertically for a superior approach to MST, avoiding striate cortex. Recording sessions lasted 4-8 hours, 5 days a week, typically collecting data from 0-5 cells per day. Each day, the monkey would perform anywhere between 500-2000 trials. The animals were given ample rest periods during each recording session. Because we were confident of our identification of area MST based on approximate location and response properties, we chose not to sacrifice the monkeys for purposes of anatomy. These monkeys went on to become subjects in subsequent investigations.

Fixation task and data collection

The animal was placed 57 cm away from a wide-field, tangent screen, projection monitor, which readily allowed stimuli as large as 40 degrees in diamter to be presented to the monkey. Trials were initiated by the appearance of a green (0.1 degree) fixation point directly ahead of the animal. The monkey was required to fixate the target and pull a lever within 600 msec of target onset. After a three second period which included the presentation of two stimuli as well as an intervening gap, the fixation point dimmed and the monkey was required to release the lever to receive a reward. Throughout the trial, eye position was monitored. If eye speeds exceeded 15 degrees/sec (as in a saccade), the trial was terminated without a reward. Software monitored eye position every 35 ms and the SD of eye position was always less than 10 min. of an arc for both animals. Data collection was controlled by a PDP-11 computer and stimulus presentation was controlled by an PC-compatible 386 computer.

Stimuli

We were interested in studying the extent to which the brain extracts the global motion of a stimulus, and accordingly we used a variety of different stimulus features to create these motion patterns.

The different visual stimuli used can be divided into different types and classes. A stimulus's class refers to whether its features are composed of random dots, lines (empty square), edges (solid square), aperture borders, or flicker. A stimulus's type refers to the motion pattern these features undergo relative to one another, namely whether they expand, rotate, contract, or spiral. An "experiment" refers to the set of data collected using the whole range of stimulus types from a single stimulus class. In an experiment, each unique stimulus was repeated 6-12 times for each unit. The trials for different classes and stimulus types were randomly interleaved. Justification for using these particular stimulus classes is given in the following section.

To understand what is meant by stimulus type it is necessary to understand the concept of a spiral space (Figure 1), originally formulated in Graziano, et al (1994). We construct this space with expansion and contraction on opposite sides of the same axis and with the two directions of rotation on opposite sides of the orthogonal axis. A stimulus whose image features have their motion vectors pointed 180 away from the center of the display (expansion) is represented straight up in this space (0 degrees); contraction is represented straight downward (180 degrees). Going from expansion to contraction is equivalent to rotating the velocity vectors of the features by 180 degrees. If, instead, these vectors are rotated 90 degrees, global rotation in either direction is obtained. For example, rotating the velocity vectors of the stimulus features 90 degrees clockwise results in clockwise rotation of the entire stimulus. Spirals result from intermediate rotations, such as the 45 degree rotations we used in these experiments.

Spirals contain elements of either expansion or contraction along with either direction of rotation, giving four basic types of spiral pattern. Using this representation, a continuous space is formed, with expansion, rotation, and contraction being discrete cardinal locations within this "spiral" space.

A stimulus "movie" is composed of 60 consecutive image frames lasting a total of 1 second. Appropriate displacement of image features between frames gives the impression of continuous motion. Six classes of stimuli were used, four of which are represented in Figure 2. The "Random Dot" class consists of dots with limited life-times (333 msec or 20 frames) that have constant, straight trajectories throughout their path. At the end of a life-time, the dot is assigned a new random location within the 20 degree diameter stimulus circle and given a trajectory and speed appropriate for its new location. The dots were relocated asynchonously, to avoid a coherent flickering of the stimulus every 333 msec. If the dot moved outside the bounds of the display window, it was immediately assigned a new, random location within the display and a new trajectory. The limited life-time of the features making up these patterns and the constant reshuffling of the dots virtually eliminates pattern artifacts. Each random-dot display contained exactly 150 dots per frame, preventing any fluctuation in overall stimulus luminance with time. For all stimulus types (patterns) the speed of each dot was a linear function of its distance from the center of the display, in this case given by the formula $S = 0.2 \times r$, where S is in (distance units)/sec and r is in distance units. During each dot life-time the velocity of the dot is held constant, eliminating any element of acceleration or curvature from its path. The direction of motion for each dot is determined by the type of global motion desired; i.e. expansion requires each dot to be moving directly away from the center of the stimulus. Because the speed of each dot is not altered by this transformation, all stimulus types have the same speed distributions.

Two other stimulus classes, Solid Square (SS) and Empty Square (ES) were created by having the corners of the squares obey the motion rules established for the random dot stimuli. However, while the dots of the RD class have limited life-times and straight paths, the corners of the squares present in the other five classes are visible for the entire movie and have acceleration and curvature consistent with their trajectories being updated every frame. The Aperture (AP) stimulus was created by moving a virtual window, identical in spatial extent and motion pattern to the squares from the ES and SS classes, over a stationary background of random dots of unlimited life-time. The background is hidden except where the square aperture window exposes the random dot background underneath. The spacing between the dots remains constant, and the dots themselves have no motion, other than to be exposed or occluded with time, depending on the motion pattern specified for the aperture.

The Flicker (FL) stimuli were identical to the SS stimuli described above, except that instead of the interior of the square being a homogeneous gray, it consisted of random pixels turning on and off every frame, creating a shimmering interior to the square. Dot density was adjusted so that the luminance contrast of the square against the background was the same as the Solid Square case. For the ES, SS, FL, and AP classes the minimum size of the square is 5 degrees of visual angle as viewed by the monkey. This occurs for the first frame of an expansion pattern and the last frame of a contraction pattern. Maximum diameter is 20 degrees.

The "Non-Fourier" stimulus was produced by creating a 20 degree square field of pixels that each have a 50% probability of being on or off. Pixel polarity does not change from frame to frame unless the imaginary border of a square obeying motion rules identical to those established for the squares described above passes over the pixel in question. Where this occurs, the polarity of the pixel reverses every frame that the virtual square border is over the pixel. In this class, motion pattern is not defined by luminance cues but by flicker in the stimulus. Note that this is not the case for the FL class, which despite containing flicker within the interior of the stimulus also possesses a square luminance border defining the motion pattern.

In the discussion that follows, an experiment refers to data recorded using a single class (RD, SS, ES, FL, AP, or NF) of stimuli for each of the eight motion types (expansion, contraction, rotation, etc.) in multiple repeats (usually around eight) of each stimulus. Figure 1 shows two superimposed tuning curves obtained by sampling either eight or sixteen directions in spiral space. As demonstrated in this figure, during preliminary experiments we determined that using eight stimulus directions gave similar response profiles as sixteen directions. We chose to sample at the lower density to save recording time. Therefore, a single experiment has around 64 trials (eight sets of eight). Sometimes less data was collected when we were unable to hold the cell or when the monkey would not cooperate with the behavior. We performed up to six different experiments on each cell. To speed data collection, in each trial two one second stimuli were presented, separated by a one second gap, during which the monkey had to continue fixating. Opposite directions of motion were presented in each trial; e.g. expansion, followed by contraction. The stimuli were all generated off-line before the experiments and displayed during the trials at a refresh rate of 60 Hz.

STIMULUS JUSTIFICATION

How different are the six stimulus classes that we are comparing? If the stimuli are too different, we would not expect a selective cell to exhibit similar responses for the different classes. On the other hand, if the stimuli are too close, the fact that they demonstrate similar tuning behavior is a trivial result. The purpose of this investigation is to test the selectivity of area MST cells for the global motion properties of a stimulus. We designed our stimulus classes using spatial features that have been investigated in previous neurophysiological work, namely random dots, aperture borders, unipolar edges, and bipolar edges. We also used both luminance cues and flicker cues to define these features. Although the fourier spectra across the five stimulus classes with luminance boundaries are quite different, they do share some similarities in their distribution of power; this is to be expected given that we intentionally kept the global motion pattern constant across classes. It may be objected that since the power spectra share similarities, the results presented below with regards to form/feature invariance are not too surprising. However, for some classes the differences are much greater than the similarities, e.g. RD vs. NF. In addition, because of non-linearities in the motion processing pathway, it may be difficult to predict unit responsiveness based purely on analysis of stimulus power spectra. For example, it has been reported that two thirds of MST neurons reverse their selectivity for translational motion direction with changes in stimulus size (Wurtz et al, 1988; Newsome et al, 1988; Komatsu et al, 1988). A cell which prefers leftward over rightward motion of a small dot may prefer the rightward motion of a large field of dots. Clearly, this could not have been predicted from the power spectra of the stimuli. Furthermore, from a theoretical perspective, the features of a stimulus which allow the nervous system to distinguish self-motion from object motion are likely related to cues that are not represented in the fourier energy spectra of the stimulus (Larish et al, 1990). If units in MST are selectively processing one of these motion sources and attempting to disregard the other, these neurons may be sensitive to such stimulus aspects. We are left with having to fall back on more descriptive comparisons of the different classes.

Early investigations in MST made preliminary observations with regards to texture and shape invariance in MST tuning (Tanaka et al, 1989; Tanaka et al, 1986). These studies showed that manipulating qualities of the stimulus such as contrast polarity (Saito et al, 1986) and texture affected the amplitude of the response to some extent but had little effect on overall selectivity for motion pattern. It was also demonstrated that, unlike the finding with regards to translation motion, sensitivity for EFCs was independent of the size of the stimulus (Duffy et al, 1991). Stimulus speed also has little influence on preferred tuning. These early studies suffered from two important limitations. At the time of these investigations, the conceptual framework of "spiral" space had yet to be introduced, and the continuity of motion pattern selectivity progressing from expansion to rotation to contraction was not understood. As a consequence, the sort of tuning curves we will construct from our data, which recover fairly precise estimates of preferred tuning direction, were not available. Previously, cells were characterized as being selective for expansion, contraction, rotation, or some combination of these motion types. As a consequence, large shifts in tuning would have been missed because of this coarse characterization. By evenly sampling spiral space with eight stimuli located 45 degrees apart in this space (8 X 45 = 360) and fitting the differential responses elicited by these stimuli to a gaussian function, preferred tuning direction can often be confidently recovered to within a few degrees. This allows us to quantitate more subtle shifts in tuning across stimulus class and stimulus location in the receptive field (Graziano et al, 1994). The second limitation of the earlier studies was the choice of stimulus classes. The range of features and cues in this investigation are much more diverse.

The random dot class is a sort of "gold standard" for work in MST and its inclusion in this study was important for comparison purposes. Random dot stimuli have the advantage of not containing any form cues. These stimuli, with some legitimacy, have often been considered as simulating optical flow patterns. Several aspects of these stimuli make them incompatible with the motion of a single object. This is evident when considering the expansion stimulus. The dots in this display move outward from the center of the pattern, consistent with the motion of an object towards an observer. However the circular boundary of this stimulus, beyond which dots are no longer visible, is stationary. Thus, the pattern remains exactly the same size. As this luminance boundary is readily visible due to the relatively high density of dots within the stimulus, the observer does not get the impression of a single approaching circular object. Instead, each dot appears as an independent feature in the visual scene. Although the dots are moving uniformly and with a motion consistent with either observer or environmental motion, they are not perceptually grouped into a single object.

A second distinguishing feature of these RD stimuli is that they do not evolve within the course of their presentation. In other words, if a random frame of the image sequence making up the RD movie were examined, it would be impossible to tell whether this frame came from the beginning or end of the sequence. Furthermore, if the instantaneous velocity fields representing the motion pattern of the image were compared at the start and end of the sequence, they also would be indistinguishable. This is because there is no element of acceleration in any part of the pattern. This is accomplished in part by using dot life-times which are finite. The effect when viewing the stimulus is that of approaching a vertical textured surface which never gets any closer to the observer. Motion and displacement are effectively dissociated. Much psychophysical evidence exists suggesting that such pure "velocity fields", despite giving rise to some ambiguities, are sufficient in many cases to allow observers to recover direction of heading (Warren et al, 1988). For these reasons, we think of this stimulus class as being "flow-like" because it captures aspects of global motion pattern sufficient for ego-motion while leaving out stimulus attributes which may be important in the perception of moving objects in the environment.

The stimuli from the other five classes do evolve with time and contain a definite form. For these classes, the edges of the squares defining the motion pattern are displaced from frame to frame. In each case there is an element of acceleration to the motion of these features. Unlike the RD class, the size of these stimuli changes over time. Viewed monocularly, the motion of the ES and SS stimuli is indistinguishable from the motion of real objects in the environment. The features move uniformly and are perceived as belonging to a single object. Only the edges of the stimuli contain any information about the motion pattern of the stimulus, because the interiors of these squares are homogeneous. Exactly the opposite is the case for the Random Dot stimulus class, where no information about motion pattern is contained in the stationary stimulus boundaries. Including both the ES and SS classes was motivated by looming detectors in other species which respond well to

SS type stimuli but poorly to ES stimuli. One such example is the locust DCMD cell (Simmons et al, 1992). Apparently, in this species there are separate channels for the representation of light/dark and dark/light transitions. At the motion pattern integration center, these two channels are mutually inhibitory and, as a consequence, the DCMD cells respond poorly to ES-type patterns, which have a close spatial juxtaposition of light/dark and dark/light luminance edges.

The time-averaged luminance boundaries for the flicker (FL) stimulus class are identical to those of the SS class. Only the interior of the stimulus has been changed. Where the SS class is a homogeneous gray, the interior of the flicker appears to "shimmer" because of individual pixels rapidly turning on and off. Studies performed in our lab comparing response properties of MT and V1 cells have shown a decreased response of units in MT to flicker compared to that of directionally selective units in V1 (Qian and Andersen, 1994). A similar trend is seen when moving up from MT to MST (Lagae et al, 1994). Neither of these studies explored what effect this suppression would have on any motion signals also present in the image. Since the signal from the luminance edges was identical for the SS and FL classes, by comparing the responses to these two classes we could get an idea of the extent of this interaction. We initially thought that such an interaction would be small as the perceptual saliency of global motion in the FL stimuli appeared similar to that of the SS and ES classes.

Similar reasoning motivated our decision to investigate the AP stimulus class. Here the luminance contrast is approximately the same as for the SS and FL classes. Like the SS class, all information contributing to motion pattern is present at the luminance boundaries of the stimulus. Unlike the SS class, which has no information regarding motion present in the interior of the stimulus, the aperture class has clearly visible stationary features that present a "no motion" signal which is in conflict with information provided by the luminance boundaries. To our knowledge, the effect of such a signal on motion processing in other areas of cortex has not been explored, although it would be interesting to investigate this with translational motion in area MT. If these stimuli drive MST cells poorly, we will have evidence that the presence of non-moving features can provide active inhibition of motion signals. Perceptually, this might seem to be the case; unlike SS, ES, and FL classes, these patterns do not produce the impression of a single object in motion.

In the previous five stimulus classes, luminance boundaries provided cues through which local motion signals are represented. In the Non-Fourier class we investigated the effect of using "second order" motion signals to define motion pattern. Traditional motion energy models cannot account for the detection of such second-order motion by the nervous system. A study by Albright (1992) showed that units in MT can respond to translational motion defined by this cue reasonably well, and we were interested to see if this was also the case in MST. It should be pointed out that these stimuli, like the AP class, have stationary features in the displays that might suppress the motion signal present in these patterns.

ANALYSIS

Extensive use of regression and hypothesis testing was made in analyzing the data. Some of these techniques are strictly valid only when linear models are considered. Because much of the time the curves we fit are nonlinear in their parameters, e.g. gaussians, the probabilities calculated are approximations. However, for large N, the various indexes used approximate actual probabilities.

For the majority of neurons, only a subset of the six experimental classes gave differential responses to different motion patterns. For the remainder, the response profiles were essentially flat. As we were interested in comparing the tuning curves obtained from the different stimulus classes, we wished to exclude those experiments where further analysis would not be productive. A crucial stage of the analysis involved fitting the data for each experiment to a gaussian function. However, we could not get the models to converge to anything reasonable when the tuning curves were close to being flat. In addition, comparing the preferred tuning of a flat response profile is meaningless. A reasonable way of excluding data would be to perform an ANOVA on each experiment class, looking at the effect of stimulus type (motion pattern) on cell firing rate. Those experiments where the variance associated with differences among stimulus type was not significant could then be excluded from analysis. The problem with this approach is that it was too sensitive, and very few experiments could be excluded at the ($p<0.05$) confidence level. The exclusion process we settled on involved regressing the data in each experiment to the horizontal line (response = constant) and then testing the hypothesis that the observed data was generated by a cell with a response profile adequately reflected by this model. We will refer to this as the flat model. This would be the appropriate model for the data in the case that stimulus type had no consistent effect on cell responsiveness. After regressing the raw data to the flat model, we obtained the residual variance. This variance is composed of two components, one part being a within stimulus type variance associated with the intrinsic variability of the data collected, obtained according to the formula:

$$s_e^2 = \frac{1}{N-n}\sum\sum(y_{ij} - \overline{y_{i\bullet}})^2,$$

where s_e^2 is an unbiased estimate of the within trial variance, N is the total number of trials from the experiment, n is the number of stimulus types, y_{ij} is the firing rate of the jth repeat of the ith stimulus type, and $\overline{y_{i\bullet}}$ is the mean firing rate for the ith stimulus type (recall that stimulus type refers to the global motion pattern of the stimulus in spiral space.) The remainder of the variance is called the "lack of fit" variance. It is equal to the total variance less the within trial variance. This value is large when determined for the flat model for cells responding preferentially to different directions of spiral motion. It represents the lack of fit in the data for the model that cannot be explained after the variance associated with randomness in cell response is factored out.

The quotient obtained by dividing this "lack of fit" variance by the within trial variance is distributed according to an F distribution with 7 and N - 8 degrees of freedom, where N is the total number of trials for the experiment (usually N is around 80). By determining where this quotients lies on the appropriate F curve, this value can be converted into a probability that is an unbiased measure of how well the data fits the proposed model. The larger this value, the better the fit. This probability measure will be referred to as the Flat Index (FI). It has a minimum value of 0 and a maximum value of 1. Note that the variance quotient is large (and the Flat Index small) when either the lack of fit is large or the within trial variance is small.

Figure 3 shows data from eight representative experiments with data sets reflecting a range of flat indexes (FI). As described below, we will use this same technique to test the goodness of fit for the gaussian models we recover for these same data sets. The FI represents the probability that the observed lack of fit from the flat model can be explained by chance. We chose to be very conservative and only excluded from further analysis those experiments where the observed lack of fit would have occurred at least 95% of the time by chance (i.e. a FI of greater than 0.95), assuming the flat model was valid. Because the data was later to be subjected to a second round of screening, we didn't wish to eliminate too many experiments at this first stage.

The experiments passing the above test were then fitted to a general gaussian function with four parameters - floor, amplitude, mean, and width according to the general formula:

$$y = a + b * \exp^{-(x-c)^2/d^2},$$

where the dependent variable, y, is firing rate, and the independent variable, x is stimulus direction. The four parameters fitted are "a" which is the floor of the gaussian function, "b" which is the amplitude, "c" which is the mean, and "d" which reflects the width. The choice of using the gaussian model to fit the data was made for two reasons. In our previous paper on MST (Graziano et al, 1994) we used the gaussian model to fit response profiles of MST cells with considerable success. As can be seen in the final frame of Figure 3, when a unit in MST gives a strong response, the profile approximates a gaussian quite well. Secondly, three of the four gaussian parameters effectively characterize relevant dimensions of a neuron's representational capacity. Amplitude of response may well be related to saliency, width to discriminability, and mean to selectivity. The statistics package Systat was used to obtain these fits, along with confidence intervals for each parameter. Mean square error was used for the loss function. Hypothesis testing was performed as for the "flat" model above, this time using the "best fit" gaussian model in place of the flat model. As before, a lack of fit index was calculated by subtracting the within trial variance from the total variance, then dividing by the within trial variance. Where this quotient fell along the appropriate F distribution recovered the probability that the observed lack of fit could have occurred by chance.

We needed to come up with an index to reflect unit differential response strength. Directional indexes that only take into account average preferred and anti-preferred responses suffer from ignoring aspects of the response profile provided by intermediate stimulus directions. Additionally, it is desirable for an index of response strength that reflects the within-type response variability of the data. The smaller this variability, the greater the representational power of a unit for a particular stimulus attribute. What was desired, in essence, was an index of "gaussianness" which would reflect both response amplitude and variability. In this way, the "robustness" of the signal could be quantitated. To do this, we needed to compare the observed data against some appropriate "flat" set of data. To obtain this, we determined the average firing rate across all trials for each individual experiment and then adjusted (by shifting) the data for each trial so that the average firing rate for each stimulus direction was identical for all eight directions sampled within a particular experiment (Figure 4). In this way, the data was "flattened." The rational for this transformation is that within trial variance remains unchanged, allowing meaningful and powerful comparisons to be made that are extensions of the hypothesis testing used above. Based on the model parameters obtained from regressing the original data to a general gaussian function, we calculated the lack of fit statistic introduced above for each experiment twice, once on the original data and once on the flattened data. From these two scores, the log of the ratio of the two probabilities was calculated. In all cases, the lack of fit of the gaussian model to the un-normalized data was appropriately small, and not significant. In most cases, the lack of fit for the flattened data was larger, particularly when the area under the model gaussian curve was large. This Gaussian Index (GI) agrees well with subjective assessments of the "gaussianness" of the data as seen in Figure 3 and is an excellent measure of response robustness. This figure shows experiments representing a range of GIs. Note that GIs are not calculated for the first two experiments; in these cases the flat index was above the threshold exclusion criteria of 0.95. We will see in the results section that a cell's flat index and the gaussian index are related empirically.

RESULTS

The basic findings of this study are evident in Figure 5. This plot shows tuning curves from a single cell where each of the six stimulus classes gave significant responses. This cell is tuned for expansion, and this property holds regardless of the features and cues used to define the motion patterns. For the Aperture class, although response to expansion remains strong, selectivity for stimulus pattern is somewhat less than for the other classes; a significant response to clockwise-rotating apertures was also recorded. Note the different scales on the axes for the six plots in this figure and that response amplitude and width does not possess the same degree of invariance as preferred direction. This unit was somewhat unusual in responding strongly to all six stimulus classes. A more representative cell is shown in Figure 6, which depicts tuning curves for a unit also tuned for expansion. Responses to FL, AP, and NF stimuli were on the order of 10 times weaker with respect to amplitude than to RD, ES, and SS. However, except for the Non-Fourier class where little selectivity is observed, a preference for expansion is maintained. This invariance with respect to feature was generally observed for all the MSTd neurons that were recorded from.

Although data from individual neurons gave strong support to the idea of form/cue invariance in MSTd, we wanted to quantitate and formalize these findings over a population of MST units. We also wanted to compare response strength across stimulus class and see how this property was related to degree of positional invariance. Because our analysis of tuning invariance depends on quantitating response robustness, we will examine this latter analysis first.

Response Strength and Experiment Screening

A total of 781 experiments were performed on these cells (639 on cells from 90-2 and 142 from 89-1). These broke down as follows: 190 Random Dot, 119 Empty Square, 184 Solid Square, 119 Flicker, 119 Aperture, 50 Non-Fourier. Many of these experiments were eliminated from further consideration due to their lack of differential responsiveness to motion pattern type, as indicated by a flat index greater than 0.95, leaving 158 (83%) Random Dot, 116 (97%) Empty Square, 152 (83%) Solid Square, 57 (48%) Flicker, 36 (30%) Aperture, 26 (52%) Non-Fourier. The percentage of ES experiments passing this test is deceptively high compared with the RD and SS classes. This is a consequence of experiments with the RD and SS patterns being the only stimulus classes investigated in 89-1. This monkey's responses in MST were not as vigorous as those in 90-2. Although we have no good explanation for this, we cannot rule out differences in visual acuity; neither monkey's vision was tested. If only 90-2's data is considered, the proportion of experiments that remained after screening via the flat index for the RD and SS classes is in line with the 83% found for the ES class. Allowing for this, the six stimulus classes can be divided into two groups, a "vigorous response" group made up of the RD, ES, and SS classes, and a "weak response" group made up of the remaining stimulus classes. This distinction is clear in examining Figure 7, which looks at the distribution of flat index by class. This same segregation will be a reoccurring theme in the analysis that follows.

Data sets from experiments with a flat index less than 0.95 were fit to gaussian curves. Using the "best-fit" parameters for a general gaussian model, we calculated the lack of fit statistic for each experiment as detailed above in Methods. The lack of fit of the gaussian model was not statistically significant for any of the data sets examined. After calculating this same statistic on the data set normalized for mean response rate ("flattening" the data) as outlined above, a gaussian index (our measure of response strength) was calculated for each experiment. If this measure of response robustness did not exceed 0.1, the experiment was discarded from further analysis. A 0.1 threshold value was chosen as it represents the point where the raw data set and the flattened data set are equally likely to have been produced by a unit whose underlying response characteristics were that of the gaussian model obtained through regression. Figure 8 looks at the GI index distribution as a function of stimulus class. The separation of the stimulus classes into two groups based on response strength that was seen for the flat index is also seen in this graph. The difference between the two stimulus groups is actually underrepresented due to the disproportionate number of experiments in the poor responding group that were screened out before this round of analysis. If the flat index screening hadn't eliminated a substantial number of the FL, AP, and NF experiments, the distribution of their GIs would have been shifted downward. Table 1 presents a summary of the screening, showing the number of experiments for the six classes passing each round of elimination.

We wanted to assess the empirical validity of our construction of the gaussian index as a measure of response robustness. An independent assessment of this property is the accuracy with which the gaussian parameters can be estimated for an experiment. If a given visual input produces a response in a particular unit that varies widely from trial to trial (high intra-type variance) the confidence with which data from such an experiment can be fitted is low, or equivalently, parameter confidence intervals large. As explained in Methods, these intervals for the four gaussian parameters estimated for each regression were calculated. Figure 9 shows the half-width of these confidence intervals for preferred tuning direction (gaussian mean) plotted against the gaussian index calculated for the same data set. A strong correlation is seen between the two scores: a high GI predicts a high level of confidence in the estimation of the fitted parameters. The confidence intervals for the other three parameters followed a similar relation (data not shown). As a general trend, higher gaussian indexes correlated with increased confidence in parameter estimation. This trend was strongest for the RD data set and weakest for the AP data set. As a general observation, the size of these confidence intervals was larger for the FL, AP, NF classes than for the other three classes, consistent with the less vigorous responses of MST units to this former group of stimuli.

Figure 10 is a scatter plot of the gaussian index against the flat index. The figure shows a strong negative correlation between the two indexes. Given either the flat index or the gaussian index for a particular experiment, the other score can be predicted with a high confidence. This suggests that the two measures are reflecting the same aspect of the data. In other words, an increase in the "non-flatness" of the data is associated with a predictable increase in the "gaussianness." This is significant for two reasons. This figure justifies our assumption that the variability of the firing rate as a function of motion type can be captured by modeling the response profile after a general gaussian function. It also justifies our first round use of the flat index to eliminate experiments from further consideration. Had we been able to fit curves for these excluded data sets, we would expect the gaussian indexes to be very low and would have excluded them from consideration based on the GI score. As can be seen from the figure, a GI of 0.1 corresponds approximately to a flat index of 0.9, justifying our higher choice of 0.95 for first round elimination.

As well as looking at population distributions, we were interested in comparing response strengths between experimental classes for individual cells. Figure 11 contains box plots showing the distribution of log ratios of the Gaussian Indexes between stimulus classes for individual units. The ratio of the GIs for the two experiments being compared is a measure of the relative robustness of the responses to the two stimulus classes. A unit in question can more precisely represent the global motion of a stimulus when the stimulus features associated with a greater GI defines the motion. The stars above the plots in Figure 11 show log ratio distributions whose means differ significantly from zero. In each graph, one of the six comparisons is between the same class, and is therefore zero in all cases, e.g. comparing RD vs. RD. This data is only shown for reference. As can be seen from Figure 11, as a population MST cells responded well to the Random Dot, Solid Square, and Empty Square classes and less well to the other three classes. Within these two groups, the responses were very similar, although often statistically different. For example, the RD class gave the most robust responses by this criteria, an effect that, although statistically significant, is small when comparing RD to the SS and ES classes. Particularly interesting is the poor response of the Flicker stimulus compared with the Solid Square, as the mean luminance contrasts and low spatial frequency structure for these two stimuli are identical. An examination of the significance of this follows in the discussion. The stimulus classes in order of increasing response strength are as follows: AP, FL, NF, ES, SS, RD.

Preferred Stimulus Pattern

Figure 12 shows the distributions of the fitted gaussian mean parameters for each stimulus class. This parameter reflects the preferred stimulus type for the unit. The length of the vector in each box corresponds to the number of units with preferred tuning direction in that range. The boxes are arranged as per the representation of "spiral space" discussed above. As has been observed in other studies of area MST, there is a predominance of cells tuned for expansion. This was true across all stimulus classes. For the Aperture class, no units tuned to CCW rotation or contraction were found and for the NF class no cells were found tuned to CW rotation. This is likely a consequence of insufficient sampling because of the small number of units that gave significant responses to these stimulus classes.

Pairwise analysis of the data from different experimental classes will emphasize comparing the means of the fitted gaussians for the different experiments done on a single cell. A total of six different stimulus classes (Random Dot, Empty Square, Solid Square, Flicker, Aperture, Non-Fourier) were used, although in few of the cells were the responses to all six classes significant enough to reach this stage for comparison. To quantify tuning invariance, we made pairwise comparisons of the gaussian means. 15 unique (30 total) potential pairwise comparisons were possible between the different classes for a single unit. These comparisons, along with the number of comparisons made are as follows: (RDvsES:105 RDvsSS:126 RDvsFL:35 RDvsAP:18 RDvsNF:19 ESvsSS:93 ESvsFL:32 ESvsAP:17 ESvsNF:19 SSvsFL:31 SSvsAP:16 SSvsNF:17 FLvsAP:10 FLvsNF:9 APvsNF:2). Table 2 shows the percentage of cases, for each comparison, where the fitted means of the classes under consideration fell outside each others 95% confidence intervals. Table 3 shows the average difference in preferred tuning between each of these stimulus classes. Clearly, those comparisons involving classes which gave poor responses tended to show larger average differences in gaussian means. Figure 13 is a series of box plots comparing the differences in these fitted means for each of the fifteen comparisons. In each case, except the comparison of AP and NF (where the N number is only 2), the difference is centered around zero. In no case was the difference between any two stimulus classes significantly different from zero (paired T-test, $p<0.05$).

In line with the idea of feature invariance, we postulated that any difference between preferred tuning directions was a consequence of the noise in the data used to fit the curves. If this was the case, experiments where the responses to the stimuli were more robust would be expected to have smaller differences between their preferred tuning directions. Figure 14 shows the magnitude of these differences as a function of the sum of the gaussian indexes of the experiments whose preferred tuning directions are being compared. As discussed above, at total of 30 (15 unique) such comparisons are possible, each of the six stimulus classes being involved in five comparisons (we are not considering comparing a stimulus class with itself, which obviously always has a difference of zero). Note that the long axis of the "wedge" shaped data is along the x-axis, indicating that the distribution is centered around zero. The variance associated with the difference in preferred tuning direction is large at small GI sums but small with high GIs. This is exactly what is expected with a stochastic distribution of the data around zero, with the GIs as a reflection of the randomness of the data. This is strong evidence for the invariance of preferred tuning direction across different stimulus classes.

Other Model Parameters

Although in studying form/cue invariance the preferred tuning direction has been given particular attention, we also examined the relative magnitudes of the other three gaussian parameters. Amplitude of response (i.e. cell firing rate in the preferred tuning direction) is often considered to be related to the perceptual saliency of the stimulus being represented. The distribution of the amplitude parameter of the fitted curves as a function of stimulus class is shown in Figure 15. Not surprisingly, this plot looks similar to Figure 8, which shows the distribution of GIs by class. Both response amplitude and GI reflect response strength, although the GI is a more sophisticated measure. As has been seen previously, the six classes of response can be divided into either strong responding or weak responding classes. A similar analysis was performed for the width of the gaussians (sigma) in Figure 16 and for the floor (estimate of firing rate in anti-preferred direction) in Figure 17. The data indicates that the width of the response curves is somewhat greater, on average, for the FL, AP, and NF classes, although this rarely reached statistical significance. However, looking at Figure 16, it is clear that the range of tuning widths is much greater for these classes. The magnitude of the floor parameter is similar across stimulus classes, except the Flicker class, where it is somewhat elevated. The reason for this is that a subpopulation of MSTd cells responded strongly to all types of motion pattern defined under the Flicker class. An example of such a unit is shown in Figure 18. Here the responses to flicker stimuli are compared to the responses for the random dot class. However, this tonic elevation in response was not observed in the majority of cases, and in some instances the opposite effect was observed — tonic inhibition. In the majority of cases, the response to the different stimulus patterns for the flicker class was generally weak, with some response selectivity for stimulus type.

A final observation has to do with the dynamics of response decay after the stimulus has been turned off. Some investigators have reported a sub-population of MSTd cells that has a particularly slow drop-off of response at the end of a trial, forming a sort decaying "tail" in the spike histograms. We observed a similar phenomenon in this investigation for some cells. Furthermore, the extent of this "tail" varied with stimulus class. Considering only those classes which gave consistently vigorous responses, ES had the slowest decay rate, followed by SS and then RD. Figure 19 shows raw spike trains and response histograms for two cells showing this effect. Although the effect was often subtle, a clear trend over many cells was evident. Also interesting was the fact that the tail was largest for stimulus patterns somewhat away from the preferred tuning direction (e.g. 45 degrees or so in spiral space).

Positional Invariance

For a few cells that we were able to hold for an extended period of time, the battery of experiments was repeated with the stimuli positioned at a different location in the unit's receptive field. Because MSTd receptive fields can be quite large, in some cases the center of the stimulus could be moved as far as 50 degrees still elicit a strong response. The property of positional invariance with respect to tuning direction has been noted in several labs, including our own (Graziano et al, 1994). In our previous investigation in MST, units were tested for this property over ranges of only about 20 degrees, due to limitations of the display device. The large screen used in this study allowed us to position the stimulus over much larger differences in visual angle. Figure 20 shows one such case where both positional and form/cue invariance were simultaneously tested in the same unit. Data was separately collected for stimulus placement in three corners of the neuron's receptive field, with these three locations forming the apexes of an equilateral triangle. Each stimulus location was 50 degrees of visual angle away from the other two. In each location, tuning curves for all six stimulus classes were obtained. To a large extent, stimulus specificity in terms of preferred response direction was maintained independent of both stimulus class and location.

DISCUSSION

This study has demonstrated form/cue invariance in macaque area MSTd, supporting the idea that this region generically represents complex motion patterns in the environment, such as those produced by moving objects. Previous studies of area MST have given only a cursory consideration to feature invariance. These studies looked at the effect of luminance boundary polarity (white stimulus on black background vs. black stimulus on white background) as well as the spatial texture of the stimulus (Saito et al, 1988). Consistent with our results, they found a relatively high degree of insensitivity to these parameters. However, these studies looked more at response amplitude than tuning direction and sampled only a few directions in spiral space. In addition, the range of feature types that we investigated in this study far exceeded those of previous work. Finally, our stimulus cues included second-order motion or "non-fourier motion" in one of the classes. This is the first time, to our knowledge, that non-fourier motion has been studied in MST.

One group (Tanaka et al, 1989ab) has obtained poor responses in area MST to stimuli smaller than 20 degrees. This would seem to make MST an unlikely candidate for the analysis of movement within the environment, as objects rarely subtend this large a visual angle. These earlier results are likely a consequence of recording under anaesthetized conditions. Working with the awake, behaving monkey, we have recorded brisk responses using stimuli as small as 5 degrees in diameter (unpublished observation). Stimuli even smaller than this may potentially drive units if the monkey is trained to attend to these patterns, such as in a discrimination task. Although not studied systematically, the presence of "rotation in depth" (rotation about the axis orthogonal to the line of sight) units reported by some workers offers further support for a motion pattern theory, as this type of motion pattern cannot be produced by self-translation in a stationary environment. In the present study, we tested this motion pattern hypothesis by generating complex motion patterns using a number of different feature types. Our choice of features and cues is discussed above in a separate section.

Neural Coding

We were interested in comparing across stimulus classes the gaussian tuning curves modeling the response profiles of MST cells tuned to complex motion patterns. Certain assumptions guided this approach, in particular that selectivity for stimulus pattern is encoded as the mean firing rate of a neuron. Although this idea is not new and has received broad support, this assumption has not received universal acceptance. Therefore, before discussing our data, we will explicitly lay out the conceptual frame-work for neural coding which motivated this study.

The four parameters of the general gaussian model describe (with regards to a particular stimulus class) a unit's preferred tuning, its amplitude of preferred response, its tuning width, and its background response. As demonstrated in this study and our previous one (Graziano et al, 1994), a continuum of response selectivity exists in MST such that all directions in "spiral space" are (Graziano et al, 1994). Because the tuning of MST cells is fairly broad, any particular stimulus will drive units whose preferred tuning lies within a range of directions around the stimulus pattern. Complex motion patterns can be distinguished by identifying the population of cells giving the strongest response to the stimulus. Therefore, motion pattern is encoded in the differential activity of separate populations of cells distinguished by their preferred tuning. In this framework, stimulus saliency is partially related to the magnitude of the response. By decreasing the motion signal in a stimulus, for example by decreasing luminance contrast, the activity of all units selective for this signal will fall off. With such changes in stimulus saliency, the location of the peak in the population response profile will remain unchanged. However, the difference between the activity of units that are tuned to the preferred stimulus compared to those tuned to the anti-preferred direction will have decreased. Accordingly the signal to noise ratio falls with luminance contrast, and precision in motion pattern representation decreases.

The third parameter of a gaussian model, tuning width (variance, or sigma), reflects the specificity of a particular unit for motion pattern. Neurons with broad tuning curves respond significantly to a wide range of motion patterns. In terms of population coding, when tuning curves are narrow fewer units are required to represent a stimulus pattern with a given level of confidence. A final consideration: If noise to added to a stimulus, depending on how the specificity of the units is constructed, the amplitude of the response in the preferred direction may remain the same, but the non-preferred responses may increase in magnitude. Like decreasing luminance contrast, the effect is again to flatten out the tuning curve, but since the floor of the response is raised, rather than the amplitude lowered, the area under the curve increases. Therefore, amplitude of preferred response is potentially a poor measure of saliency.

In this study, we recovered multiple tuning curves from the same unit, but used different stimulus attributes to define the motion pattern for each of these curves. Systematic differences in these curves between stimulus classes presumably reflect variations in how the motion signal is represented. Our principle prediction, based on the model for neural coding presented above, was that if MST is generically representing motion pattern, preferred stimulus type should remain invariant across stimulus class. Secondly, we would expect to see some correlation of response strength with the saliency of the perception.

Tuning Invariance

Our experiments indicate that the preferred motion pattern for a majority of cells in area MST is not dependent on the features that define the motion. When statistically significant differences in preferred tuning existed, the magnitude of these differences tended to be small, compared with the possible range of preferred tunings. As a population, the more confident we were with the gaussian models we obtained from the data, the smaller these differences were. Other aspects of the tuning curves, i.e. response height, did vary depending on the features used to define the motion type of the stimulus. It would be reasonable to suggest that these differences reflected differences in stimulus saliency. From our data, it is evident that the stimulus classes RD, ES, and SS gave more robust responses than FL, AP, and NF. Subjectively, the strength of the motion associated with RD, ES, and SS is much greater than with AP and NF, although a future study needs to be done to quantitate this. The poor response to FL is somewhat puzzling, given the apparently strong motion signal associated with this stimulus.

The SS and FL stimuli are identical in terms of their time-averaged luminance contrast. The only difference is that whereas the square in SS is a homogenous gray, FL contains a square composed of flickering dots. The tuning curves obtained from the SS class were, on the whole, much more robust than those obtained using the FL class. A "motion opponency" stage in the processing of motion signals may explain this result (Qian and Andersen, 1994). In a natural scene, motion borders have a large amount of random energy on top of the "true" motion energy associated with movement of the border. This is the case whenever a motion boundary is occluding or uncovering background features. This random spatio-temporal energy is viewed as noise by a motion detection system. Much of this noise can be filtered out by a process of "motion opponency", whereby motion signals are smoothed in a local region of image space. The 3-D fourier spectrum of the flicker stimulus possesses a considerable background of energy in random directions as well as a "true" motion signal present at the luminance edges. Apparently, for the case of the FL class, the noise signal is overwhelming the motion signal and, after the opponency stage, very little signal is left for MST cells to detect. The opponency stage has been proposed, by other workers in our lab, to be located at area MT, although it may be occurring in stages from V1 up through MST. Based on a motion-opponency model, Qian and Andersen (1994) predicted for area MT that the greater a cell's direction selectivity, the stronger the motion opponency, and the weaker the cell's response to flicker. This same prediction should hold for MST cells, whose major inputs are from MT units. We have yet to analyze our data to see if a cell's response to the Flicker stimuli is correlated in any way with motion pattern specificity.

Although this explanation is reasonable, it doesn't account for our perception of the FL stimuli. Casual observations comparing SS and FL classes suggest similar perceptual saliency to their motion patterns. If MST processes motion pattern and a stimulus class fails to elicit a tuned response in this region, we should have difficulty perceiving motion patterns defined with these features. To resolve this difficulty, we postulate an attentional mechanism. Since the monkey is simply trained to fixate straight ahead during the experiment, he is unlikely to be attending to the patterns which are presented. This is not true of human observers who are examining these stimuli. We predict that if the monkey was trained to discriminate motion pattern during the recording paradigm, that responses to these Flicker patterns would rival those of the Solid Square. If this turns out not to be the case, this would be a convincing blow to the motion pattern hypothesis.

Weak responses were also observed for the Non-Fourier patterns. Although there are some differences in the way the stimuli were designed, studies with non-fourier translational motion in MT showed much more vigorous responses. Albright (1992) reported that the response in MT to non-fourier patterns was approximately half that obtained with luminance defined motion. This difference may be related to the increased suppression of flicker in MST compared with MT (Lagae et al, 1994). It is likely that much of the second-order motion signal is lost in the process of spatio-temporal smoothing that accompanies noise suppression.

For the case of the Aperture class, given that the luminance contrast borders of these stimuli were similar to those of the ES and SS classes, the fact that MST responds poorly to this stimulus class is interesting. The presence of stationary features (in this case random dots) in the stimulus may be actively suppressing the motion signals provided from the square luminance borders. The interiors of the SS and ES classes contain no motion signals at all, because of an absence of texture. Using these two classes, a strongly tuned response to was obtained, presumably based on the differential motion of the four luminance edges. We suggest that the weak responses elicited with the aperture class argues for MST having a role in perception of object motion and against this region being involved in the processing of ego-motion. Consider the case of an observer moving forward across a field, with a mountain range in the distance. With his eyes fixated on the horizon, these far off features have essentially no retinal motion. However, features more peripheral, such as the ground over which he is walking, define motion vectors consistent with forward locomotion. Therefore, in developing an egomotion system it would be important that null motion signals from stationary retinal features would not suppress responses from other motion sources. On the other hand, we would expect an object motion system to have opposite characteristics. The fact that the velocity field varies continuously in regions defining objects is important in scene segmentation. The aperture stimuli used in this investigation could not correspond to any real world object because the motion defined by the luminance edges of the square is not matched by any motion of the internal features. For a system representing object motion, this inconsistency would be important to detect. The poor responses to the Aperture class are consistent with the nervous system being sensitive to such stimulus qualities. Perceptually, the Aperture class does not give the impression of a moving object, unlike the SS and ES classes. Therefore, there is agreement between the perception and the relative activity in MST.

Unfortunately, the Random Dot class presents a challenge to this formulation. The internal features of these stimuli are consistent with rigid object motion, while the edges (defined by a luminance border) are stationary. For the case of rotation this is not a problem, because the circular border of the stimulus cannot supply a motion signal. With normal object expansion and contraction the stimulus should change size, unlike ther Random Dot patterns. Our perceptions are in agreement with this analysis: while the rotating random dot pattern looks like a single, rigid object, the expanding and contracting patterns do not give the impression of an approaching or receding object. The fact that MST responds well to this stimulus class is difficult to explain, assuming a selectivity for object motion.

Although the correlation of saliency with response amplitude is unclear, whenever robust tuning curves could be established for the same unit across multiple stimulus classes, the preferred tuning direction was similar. Apparently, in terms of stimulus selectivity, MST abstracts information related to the spatio-temporal pattern of the stimulus, ignoring other attributes. In our lab's previous paper on MST, we introduced an analogy with the inferotemporal region (area IT) to express this idea (Graziano et al, 1994). Area IT has long been thought to represent the spatial organization of stimulus features. In this region, investigators have reported the existence of "face cells", "toilet brush cells", and the like (Gross et al, 1972). We have discussed evidence that a specificity for motion pattern has exists in MST. The outputs of IT and MST converge in parietal cortex, perhaps to pool together information from these different processing streams. The information represented in these two pathways is, to some extent, orthogonal, as we have demonstrated for motion pattern in this study. This notion was well expressed in (Stoner et al, 1992) with reference to the representation of translational motion: "The direction and rate at which an object moves are not normally correlated with the manifold physical cues (for example, brightness, and texture) that enable it to be seen. As befits its goals, human perception of visual motion largely evades this diversity of cues for image form; direction and rate of motion are perceived in a fashion that does not depend on the physical characteristics of the object." This property of processing systems in the brain is referred to a form/cue invariance. The reason this quality has emerged repeatedly in the brain is because of the computational efficiency is affords. As an analogy, consider our system of mathematics. We have one system to manipulate any kind of quantity, whether it be number of dogs, birds, or golf balls. The numerical computations we perform on these things do not depend on what the numbers are representing. Imagine having to use a different mathematical set of rules for every type of thing in the world. By breaking any analysis down into orthogonally separate parts, we can avoid elements of redundancy. If the brain had a separate system for the motion analysis of squares, circles, triangles, and what not, the brain would be prohibitively large. Thus, motion analysis occurs independently of the cues/features used to define this motion. Knowing that the brain exploits this strategy, we can design experiments that attempt to discover which axis of the stimulus "attribute space" a region is sensitive to. By finding axes along which no differential response is detected, we further narrow our understanding of which stimulus characteristics are being represented.

This notion of attribute orthogonality can be thought of in terms of "labeled lines." In this conception, perceptual systems are segmented into distinct processing channels. Each channel can be considered a separate axis in a multidimensional attribute space. This space is defined prior to the introduction of any particular stimulus. Every possible stimulus that we can perceive falls somewhere in this space. It is the combination of activities in each of these channels which determines where in the space the stimulus is represented and what is perceived. By having these axes as independent (orthogonal) as possible, we not only reduce redundancy, but prevent interference between different processing domains which need to remain distinct.

These concepts have received attention in the psychophysical literature (Stoner et al, 1992; Berkeley et al, 1994). Single unit recording studies in area V1 of the macaque have demonstrated that these neurons are form/cue invariant in their selectivity for image contours (Albright et al, 1989) and in MT for local linear motion signals (Albright, 1992). This independence holds for bandwidth and preferred tuning, but not for response amplitude. The average difference between stimulus classes in preferred tuning direction did not vary from zero, and no more than 30 percent of the differences were outside the range of +/- 45 degrees. This is similar to our findings in MST for motion pattern. These authors suggested, as we have, that amplitude is related to attribute saliency. Cue invariance has been reported in area IT as well (Sary et al, 1993). These investigators recognized that the perception of shape is invariant with respect to location in space, size, and the cues which define the shape. Using cues based on differences in luminance, motion, and texture, they found in IT a physiological correlate of this perceptual invariance. The neurons ignored aspects of the stimulus unrelated to spatial structure.

We believe the results of this investigation favor a direct role for area MSTd in the processing of object motion. The available evidence cannot rule out a role for this region in processing "optical flow" related information. Rather, it is likely that MST is a least part of the processing stream analyzing optical flow. As such, it is prudent at this juncture to characterize MST as a generic "pattern motion" integration center.

REFERENCES

Albright TD (1984) Direction and Orientation Selectivity of Neurons in Visual Area MT of the Macaque. Journal of Neurophysiology 52(6): 1106-1130.

Albright TD (1992) Form-Cue Invariant Motion Processing in Primate Visual Cortex. Science 28: 1141-1143.

Albright TD, Chaudhuri A (1989) Society for Neuroscience Abstracts 15: 323.

Andersen GJ (1986) Perception of Self-Motion: Psychophysical and Computational Approaches. Psychological Bulletin 99(1): 52-65.

Andersen RA (1989) Visual and eye movement functions of the posterior parietal cortex. Annual Review of Neuroscience 12:337-403.

Berkley MA, Debruyn B, Orban G (1994) Illusory, motion, and luminance-defined contours interact in the human visual system. Vision Research 34(2): 209-16

Braddick OJ, Holliday IE (1991) Serial search for targets defined by divergence or deformation of optic flow. Perception 31:345-354.

Desimone R, Albright TD, Gross CG, Bruce CJ (1984) Stimulus selective properties of inferior temporal neurons in the macaque. Journal of Neuroscience 4:2051-2062.

Desimone R, Ungerleider LG (1986) Multiple Visual Areas in the Caudal Superior Temporal Sulcus of the Macaque. The Journal of Comparative Neurology 248: 164-189.

Duffy CJ, Wurtz RH (1991a) Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. Journal of Neurophysiology 65(6): 1329-1345.

Duffy CJ, Wurtz RH (1991b) Sensitivity of MST neurons to optic flow stimuli. II. Mechanisms of response selectivity revealed by small-field stimuli. Journal of Neurophysiology 65(6): 1346-1359.

Duffy CJ, Wurtz RH (1993) An illusory transformation of optic flow fields. Vision Research 33(11): 1481-90.

Egelhaaf M, Hausen K, Reichardt W, Wehrhahn C (1988) Visual course control in flies relies on neuronal computation of object and background motion. Trends in Neuroscience 11(8): 351-358.

Erickson RG, Thier P (1991) A neuronal correlate of spatial stability during periods of self-induced visual motion. Experimental Brain Research 86: 608-616.

Freeman TC, Harris HG (1992) Human sensitivity to expanding and rotating motion: Effects of complementary masking and directional structure. Vision Research 32: 81-87.

Gibson JJ (1950) The perception of the visual world. Boston: Houghton Mifflin.

Graziano MSA, Andersen RA, Snowden RJ (1994) Tuning of MST Neurons to Spiral Motions. Journal of Neuroscience 14(1):54-67.

Gross CG, Rocha-Miranda CE, Bender DB (1972) Visual properties of neurons in inferotempral cortex of the macaque. Journal of Physiology (London) 35:96-111.

Hildreth EC (1991) Recovering heading for visually guided navigation. Vision Research 32(6): 1177-1192

Hildreth EC, Koch C (1987) The Analysis of Motion: From Computational Theory to Neuronal Mechanism. Annual Review of Neuroscience 10: 477-533.

Judge SJ (1990) Vision. Knowing where you're going [news; comments]. Nature 348(6297).

Koenderink JJ, van Doorn AJ (1981) Exterospecific component of the motion parallax field. Journal Optical Society of America 71(8): 953-957.

Komatsu H, Wurtz RH (1988) Relation of cortical areas MT and MST to pursuit eye movements. III. Interaction with full-field visual stimulation. Journal of Neurophysiology 60(2): 621-644.

Kruk R, Regan D (1983) Visual test results compared with flying performance in telemetry-tracked aircraft. Aviation, Space, and Environmental Medicine 54:906-911.

Lagae L, Maes H, Raiguel S, Xiao DK & Orban, GA (1994). Responses of macaque STS neurons to optic flow components: a comparison of areas MT and MST. Journal of Neurophysiology 71:1597-626.

Larish JF, Flach JM (1990) Sources of Optical Information Useful for Perception of Speed of Rectilinear Self-Motion. Journal of Experimental Psychology: Human Perception and Performance 16(2): 295-302.

Lisberger SG, Morris EJ, and Tychsen, L (1987) Visual motion processing and sensory-motor integration for smooth-pursuit eye movements. Annual Review of Neuroscience 10:97-129.

Maioli MG, Squartrito S, Domeniconi R (1989) Projections from visual cortical areas of the superior temporal sulcus to the lateral terminal nucleus of the accessory optic system in macque monkeys. Brain Research 498: 389-392.

Maunsell J, van Essen DC (1983a) Fuctional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. Journal of Neurophysiology 49(5): 1127-1147.

Maunsell J, van Essen DC (1983b) Functional properities of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitiviy to binoncular disparity. Journal of Neurophysiology 49(5): 1148-1167.

Nakayama K (1985) Biological Image Motion Processing: A Review. Vision Research 25(5): 625-660.

Newsome WT, Wurtz RH, Komatsu H (1988) Relation of Cortical areas MT and MST to pursuit eye movements. II. Differentiation of retinal from extraretinal inputs. Journal of Neurophysiology 60(2): 604-620.

Orban GA, Lagae L, Verri A, Raiguel S, Xiao D, Maes H, Torre V (1992) First-Order analysis of optical flow in monkey brain. Proceedings of the National Academy of Science 89: 2595-2599.

Poggio T, Verri A, Torre V (1991) Greens Theorems and Qualitative Properties of the Optical Flow (1991) MIT AI Memo 1289.

Prazdny K (1980) Egomotion and relative depth map from optical flow. Biological Cybernetics 36: 87-102.

Probst T, Brandt T, Degner D (1986) Object-motion detection affected by concurrent self-motion perception: psychophysics of a new phenomenon. Behavioral Brain Research 22(1): 1-11.

Regan D (1986) Visual processing of four kinds of relative motion. Vision Research 26:127-145.

Regan D, Beverley KI (1978) Looming detectors in the human visual pathway. Vision Research 18: 415-421.

Regan D, Beverley, KI (1982) How Do We Avoid Confounding the Direction We are Looking and the Direction We are Going. Science 215: 194-196.

Rieger JH, Lawton DT (1985) Processing differential image motion. Journal Optical Society of America 2(2): 354-359.

Roy JP, Hidehiko K, Wurtz R (1992) Disparity Sensitivity of Neurons in Monkey Extrastriate Area MST. Journal of Neuroscience 12(7): 2478-2492.

Royden CS, Banks MS, Crowell JA (1992) The perception of heading during eye movements. Nature 360(6404): 583-5.

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(1): 145-157.

Sakata H, Shibutani H, Kawano K, Harrington TL (1985) Neural mechanisms of space vision in the parietal association cortex of the monkey. Vision Research 25:453-463.

Sary G, Vogels R, Orban GA (1993) Cue-invariant shape selectivity of macaque inferior temporal neurons. Science 260(5110): 995-7.

Schwartz EL, Desimone R, Albright T, Gross C (1983) Shape recognition and inferior parietal neurons. Proceedings of the National Academy of Science USA 80: 5776-5778.

Sekuler AB (1992) Simple-pooling of unidirectional motion predicts speed discrimination for looming stimuli. Vision Research 32(12): 2277-88.

Shiffrar M, Pavel M (1991) Percepts of Rigid Motion Within and Across Apertures. Journal of Experimental Psychology 17(3): 749-761.

Simmons PJ, Rind FC (1992) Orthopteran DCMD neuron: a reevaluation of responses to moving objects. II. Critical cues for detecting approaching objects. Journal of Neurophysiology 68(5): 1667-82.

Stoner GR, Albright TD (1992) Motion Coherency Rules are Form-Cue Invariant. Vision Research 32(3): 465-475.

Straube A, Paulus W, Brandt T (1990) Influence of visual blur on object-motion detection, self-motion detection and postural balance. Behavioural Brain Research 40: 1-6.

Swanston MT, Wade NJ, Day RH (1987) The representation of uniform motion in vision. Perception 16(2): 143-59.

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(1): 134-144.

Tanaka K, Fukada Y, Saito H (1989) Underlying Mechanisms of the Response Specificity of Expansion/Contraction and Rotation Cells in the Dorsal Part of the Medial Superior Temporal Area of the Macaque Monkey. Journal of Neurophysiology 42(3): 642-656.

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(3): 626-641.

Turano K, Wang X (1994) Visual discrimination between a curved and straight path of self motion: effects of forward speed. Vision Research 34(1): 107-14.

Ungerleider LG Desimone R (1986a) Projections to the superior temporal sulcus from the central and peripheral field representations of V1 and V2. Journal of Comparative Neurology 248:147-163.

Ungerleider LG, Desimone R (1986b) Cortical Connections of Visual Area MT in the Macaque. Journal of Comparative Neurology 248: 190-222.

van den Berg, AV (1992) Robustness of perception of heading from optic flow. Vision Research 32(7): 1285-1296.

van den Berg AV (1993) Perception of heading [letter]. Nature 365(6446): 497-8.

van Essen DC, Maunsell JHR (1983) Hierarchical organization and functional streams in the visual cortex. Trends in Neuroscience 6: 370-375.

Warren WH, Blackwell AW, Kurtz KJ, Hatsopoulos NG, Kalish ML (1991) On the sufficiency of the velocity field for perception of heading. Biological Cybernetics 65: 311-320.

Warren WH, Hannon DJ (1988) Direction of self-motion is perceived from optical flow. Nature 10(336): 162-163.

Warren WH, Hannon DJ (1990) Eye movements and optical flow. Journal Optical Society of America 7(1): 160-169.

Warren WH, Mestre DR, Blackwell AW, Morris MW (1991) Perception of circular heading from optical flow. Journal of Experimental Psychology: Human Perception and Performance 17(1): 28-43.

Werkhoven P, Koenderink JJ (1991) Visual processing of rotary motion. Perception and Psychophysics 49: 73-82.

Wertheim, AH (1981) On the relativity of perceived motion. Acta Psychologie. 48: 97-110.

Wurtz RH, Komatsu H (1988) Relation of cortical area MT and MST to pursuit eye movements. I. Localization and visual properties of neurons. Journal of Neurophysiology 60(2): 580-603.

Wylie DR, Frost BJ (1991) Purkinje cells in the vestibulocerbellum of the pigeon respond best to either translational or rotational wholefield visual motion. Experimental Brain Research 86(1): 229-232.

Wylie DR, Frost BJ (1993) Responses of pigeon vestibulocerebellar neurons to optokinetic stimulation. II. The 3-dimensional reference frame of rotation neurons in the flocculus. Journal of Neurophysiology 70(6): 2647-59.

Wylie DR, Kripalani T, Frost BJ (1993) Responses of pigeon vestibulocerebellar neurons to optokinetic stimulation. I. Functional organization of neurons discriminating between translational and rotational visual flow. Journal of Neurophysiology 70(6): 2632-46.

LEGENDS

Table 1: Number of experiments analyzed for each stimulus class at each round of elimination. Numbers in the first column are the total number of experiments that data was collected for. Experiments were excluded from further consideration at two sequential stages. At the first stage, only those experiments with a Flat Index (FI) < 0.95 were regressed to a gaussian model with subsequent calculation of a gaussian index (GI) for that data. If an experiment's GI > 0.1, it was used in future comparisons. The right two columns give averages of the flat index and gaussian index for all experiments in which they were calculated. The GIs fro FL, AP, and NF classes would have been even lower had not a substantial number of these experiments been eliminated at the previous round. RD = Random Dot; ES = Empty Square, SS = Solid Square, FL = flicker, AP = aperture; NF = Non-Fourier, AP = Aperture, FI = Flat Index, GI = Gaussian Index.

Table 2: Percentages of preferred tuning directions statistically different between stimulus classes. A particular comparison is represented by the intersection of a row and a column labeled with the classes being compared. Numbers in the cells are the percentage of instances where the fitted means of two experiment's tuning curves fell outside each others 95% confidence intervals obtained during regression. Only experiments where both the GI exceeded 0.1 were used for this comparison. RD = Random Dot; ES = Empty Square, SS = Solid Square, FL = flicker, AP = aperture; NF = Non-Fourier, AP = Aperture.

Table 3. Average stimulus class differences in preferred tuning. This table is in the same format as Table 2. Numbers in boxes formed by the intersection of a columns and a row represent the average observed difference in preferred tuning direction between stimulus classes for individual units. Numbers are all positive because absolute values of these differences were taken. If feature invariance did not occur in MST, these averages would all be distributed around 180 degrees. Thus, a considerable degree of invariance is indicated. Note that numbers are smaller when comparisons are made between classes that generally provided strong response profiles (RD, ES, SS).

Figure 1: Spiral space explained. In this representation, expansion/contraction are on opposite sides of the vertical axis, while the two directions of rotation are on opposite sides of the orthogonal axis. Intermediate orientations between these cardinal axes represent spiral patterns. Expansion (straight up) is assigned an angular value of 0 degrees by convention, with the angles increasing as one moves clockwise in this space. Distance from the origin is a measure of firing rate; angle reflects the type of motion pattern. Data from the two curves displayed on this polar plot was obtained from the same cell using the same stimulus class (Random Dot). With one curve we sampled the space every 22.5 degrees (sixteen points), the other curve every 45 degrees (eight points). Each point plotted represents the average firing rate for

that stimulus type (motion pattern) pooled from repeated, randomly interleaved trials. Note the similarity of the two curves obtained at the two sampling frequencies. The lines emanating from the origin represent the preferred tuning directions recovered from this data after regression to gaussians. This particular unit was tuned to a spiral containing elements of both expansion and clockwise rotation. The similarity of tuning curves for the two sampling densities motivated us to use the lower sampling frequency so that more data could be collected per recording day.

Figure 2: Four of this six stimulus types that we used for our investigation. **A** shows representative motion vectors for a spiral stimulus that is midway between expansion and counter-clockwise rotation (corresponding to an angular location of 315 degrees in spiral space). Note that the path of these dots has no curvature, the local motion of the dots being linear. The lengths of these vectors increases with distance from the center of the stimulus. In **B**, **C**, and **D** we show representative motion pattern stimuli at two points in time. As the squares represented are both rotating and getting larger with time, their motion pattern is also that of a spiral. In the Aperture class, the texture elements making up the interior of the square do not move with the edges of this stimulus. The spacing and placement of these texture elements remains unchanged. For the square stimuli, maximum spatial extent is 20 degrees, minimum spatial extent 5 degrees. Although the patterns are all represented as black against a white background, on the screen the polarity was actually opposite. The luminance contrast of the Aperture and Solid Square classes were actually more similar than they appear in this figure. Two additional classes, Flicker and Non-Fourier were also used in this investigation, but the nature of these stimuli prevented a convenient static representation.

Figure 3: Raw data demonstrating various Flat and Gaussian indexes. Moving across and then down, the data increasingly takes on a more gaussian shaped profile. The variance within stimulus types also decreases. No GI was obtained for the first data set because the Flat Index exceeded 0.95. The second frame shows data from an experiment near our threshold criteria for exclusion based on the Flat Index. Multiple data points within a particular graph for the same tuning direction represent repeated trials. The horizontal axis label, "NORMDIR", stands for "normalized direction" and represents the direction in spiral space of the stimulus that gave the response indicated by the ordinate value. Before curves could be fit to the data, the data was shifted so that the peak of the gaussian was centered at approximately 180 degrees. This facilitated regression in an analysis package, "Systat", that did not provide for circular statistics. Therefore, it was not the case that the units in the six experiments represented in this figure all had there preferred tuning direction near 180 degrees (contraction). After regression, the parameters recovered were shifted in the opposite direction by an equivalent amount.

Figure 4: Process of Normalization. A) The upper frame shows the raw data from one experiment (i.e. one stimulus class, multiple motion types). This graph is identical in form with those described above in Figure 3. The horizontal axis represents the motion pattern of the inducing stimulus, in angular units of spiral space. The ordinate represents the magnitude of that response (in spikes/sec) to the stimulus. B) The bottom frame shows the consequence of "flattening" the data, as outlined in the text. The within trial variance as well as the mean firing rate with respect to the entire data set remains constant, but the average firing rate is now the same across all stimulus directions. By statistically comparing the top and bottom data sets we achieved as measure of "gaussianness."

Figure 5: Comparison of tuning curves from a single unit for the six stimulus classes. This comes from one of the few cells (B10800) where the Flat Index and Gaussian Index criterion were both met for all six experimental classes. The location of each data point in these polar plots reflects both the magnitude of the response and the stimulus type used to elicit the response. Distance away from the origin indicates response strength in unit spikes/second while the angle is a function of the stimulus motion pattern inducing this activity. This particular unit is tuned for expansion. Note the similarity in the curves in terms of both orientation and shape, although there is some variation with the aperture class. In this case, tuning specificity is less pronounced and this unit responds fairly strongly to apertures containing clockwise rotational movement (equivalent to 90 degrees in spiral space.)

Figure 6. Same information present in Figure 5, but for unit B08200. This unit is also tuned for expansion. Note that the scale of the radius axis is different for each of the plots. Accordingly, responses to the FL, AP, and NF classes are on the order of ten times weaker than those to the RD, SS, and ES classes. The response to the Non-Fourier stimuli was not well tuned at all; little selectivity in activity is demonstrated. This is likely a function of the poor signal to noise ratio associated with the low firing rates obtained from these stimuli. This general behavior was typical of the units that we found.

Figure 7: Flat Index distribution by class. This is a box plot showing the distribution of Flat Indexes divided according to stimulus class. Each of the six plots should be viewed as a sort of compact histogram with a gaussian shaped distribution. The solid black squares indicate the means of this distribution. The short horizontal lines bisecting the vertical rectangles represent the median of the distribution. The vertical rectangle includes 75% of all index scores. The "whiskers" attached to both ends of these rectangles extend out to include 90% of the data. The six stimulus classes fall fairly neatly into two different groups. The FL, AP, and NF classes, with high Flat Index scores, are placed in a "poor responding" group, while the remaining three classes generally elicited more vigorous responses as reflected in their lower LI distributions. RD = Random Dot; ES = Empty Square, SS = Solid Square, FL = flicker, AP = aperture; NF = Non-Fourier, AP = Aperture, FI = Flat Index.

Figure 8. Gaussian Index distribution by class. This box plot shows the distribution of gaussian indexes divided according to stimulus class. The notation of the graph is identical to that of Figure 7 above. Each plot should be viewed as a sort of compact histogram providing information about the general form of the distribution. The solid black squares indicate the means for each distribution. The short horizontal lines bisecting the vertical rectangles represent the medians. The vertical rectangle includes 75% of all index scores. The "whiskers" attached to both ends of these rectangles extend out to include 90% of the data. The six stimulus classes fall fairly neatly into two different groups. The FL, AP, and NF classes, with low Gaussian Index scores, are placed into the "poor responding" group while the remaining three classes establish a "strong responding" category. RD = Random Dot; ES = Empty Square, SS = Solid Square, FL = flicker, AP = aperture; NF = Non-Fourier, GI = Gaussian Index.

Figure 9: Uncertainty in estimated preferred tuning direction versus gaussian index. 95% confidence intervals were obtained for the means (preferred response direction) obtained for the general gaussian function fit to each experimental data set. If the Gaussian Index reflects the variability in the data, it would be expected that there would be an inverse correlation between this index and the uncertainty with which preferred tuning direction can be recovered. These scatter plots of the half-width of the 95% confidence intervals against Gaussian Index demonstrate that such a relation holds. High indexes are associated with small confidence intervals, indicating low levels of uncertainty with regards to parameter estimation.

Figure 10: Scatter Plot of Gaussian Index vs. Flat Index divided into experimental classes. The distribution of each plot suggests an intimate relationship between the two measures. Given one index for a particular experiment, the other index can be predicted with a high degree of certainty. This association increases our confidence that the Gaussian Index is an adequate reflection of the differential responsiveness of the data set.

Figure 11: Box plots showing the distribution of the log ratios of Gaussian Indexes between experiments done on the same unit with different stimulus classes. A value of zero indicates that the GI's for the data from the two experiments were equal. This was obviously the case when a data set was compared against itself; these cases are shown only as a point of reference. A positive value indicates that exp1 of the exp1/exp2 pair shown on the x axis had the higher Gaussian Index. Accordingly, from this data it is clear that the RD class, on average, gave stronger responses than any other experimental group. Note also that the SS, ES, and RD stimuli consistently sported higher GIs than the NF, AP, and FL stimuli. Stars above a plot indicate that the mean of this distribution was significantly different from zero (paired TTest, $p<0.05$).

Figure 12: Distribution of preferred tuning directions (gaussian means) by stimulus class. These distributions were not evenly spread out over

spiral space. An over representation of units tuned to expansion was observed. A second, smaller peak, for contraction also is evident. For purposes of this analysis, spiral space was divided into 8 equally sized wedged pieces, as shown in each of the six plots. These diagrams are laid out in a similar fashion to the polar plots in preceding figures. A response profile was characterized as being centered around "expansion", for example, if the preferred tuning recovered for this experiment was between -22.5 degrees (same as 337.5 degrees) and 22.5 degrees. Each box, representing one of the eight direction types, contains a vector whose orientation points in the direction of spiral space being considered, and the length of the vector reflects the number of units with this tuning preference. A similar distribution is seen for each stimulus class. The inconsistencies with regards to the FL, AP, and NF classes are likely a consequence of the small numbers of units for which tuning curves were obtained.

Figure 13: Box Plots showing the distribution of the differences between the fitted means (preferred tuning directions) for curves obtained using different stimulus classes. In this case, the sign of these differences has been preserved. The x axis shows the two data sets being compared, the y axis reflects the distributions of these differences. In the convention we have adopted, a positive differences indicates that the shift in preferred tuning from exp_1 to exp_2 was clockwise, a negative difference counter-clockwise. The filled square represents the population mean for this shift; the top, bottom, and line through the middle of the box correspond to the 75th percentile (top quartile), 25th percentile (bottom quartile), and 50th percentile (median) respectively. The whiskers on the bottom extend from the 10th percentile (bottom decile) and 90th percentile (top decile). The comparison of one experimental type against itself is obviously zero and this data is shown only for reference.

Figure 14: Differences in preferred tuning direction as a function of Gaussian Index. Based on the hypothesis that MST selectivity is invariant with regards to the form, feature, and class defining motion pattern, we predicted that any variation in preferred tuning between classes was largely a function of the inherent variability in the data. Because gaussian index reflects response robustness, we predicted that these tuning differences would be smaller when calculated between experiments with high GIs. In this figure, by plotting these differences in preferred tuning against the sum of the GIs of the experiments compared, it can be seen that this is the case. For each plot, five types of potential comparisons are made, with the title of the plot indicating the stimulus class in common for each of these comparisons. For example, the graph at the upper left makes comparisons RD-ES, RD-SS, RD-FL, RD-AP, and RD-NF.

Figure 15: This box plot shows distributions, broken down according to stimulus class, for the gaussian amplitude parameters recovered from the experimental data sets. The amplitude is a measure of the response strength of a unit to its preferred stimulus pattern. The RD, ES, and SS

classes consistently demonstrated greater response amplitudes than the other three classes. This diagram actually underrepresents this trend because many nearly flat sets of data obtained from FL, AP, and NF classes was eliminated before this round of analysis through a high Flat Index score. RD AMP = Random Dot amplitude, ES AMP = Empty Square amplitude, SS AMP = Solid Square amplitude, FL AMP = Flicker amplitude, AP AMP = Aperture amplitude, NF AMP = Non-Fourier amplitude.

Figure 16. This box plot shows distributions, broken down into stimulus class, for the gaussian parameter sigma (width, or variance) recovered from experimental data sets. The variance is a measure of the response selectivity of a unit to its preferred stimulus pattern. Higher values for this parameter describe broader tuning curves. The distribution of this parameter was somewhat wider for the FL, AP, and NF classes, but the means of these distributions were similar.

Figure 17: This box plot shows distributions, broken down into stimulus class, for the gaussian floor parameters. The floor is a measure of the response of a unit to its anti-preferred stimulus pattern and is often close to the baseline firing rate of the neuron. The distribution of this parameter was somewhat wider for the FL, AP, and NF classes, but the means of these distributions were similar. An explanation for the slight upward shift in the distribution of the FL class is described in the text.

Figure 18: Responses of MST units to the Flicker stimulus class. For a subclass of neurons in this region, the response to the FL stimulus class was always well above the background firing rate of the cell, even for the anti-preferred tuning direction. A) Raw spike data and spike histograms for unit B11000. Pictures above the data represent the stimulus pattern which elicited the responses underneath. The responses to the RD and FL classes are compared. For the FL class, note that both responses in the preferred and anti-preferred directions are shifted upward. B) Polar plot of this data, comparing the tuning curves.

Figure 19: Three cells demonstrating the response "tail" following the extinguishing of the stimulus. Vertical lines in the plots indicate stimulus on and off times, as labeled in the figure. The response decay constant is fairly large for these cells, particularly with regards to the ES class. This aspect of the response was true for about half of the cells investigated. Interestingly, the effect appears strongest in the responses to patterns somewhat away from the preferred tuning direction of the unit.

Figure 20: Three sets of polar plots showing tuning response invariance with respect to both stimulus feature and location within the receptive field. This data was recorded from a cell (B11300) with a particularly large receptive field. This allowed for placement of the stimuli at three spatial locations, each 50 degrees apart from one another, forming the corners of an equilateral triangle. Data from each polar plot was obtained from a different spatial location. Although response amplitude

varied considerably over feature classes and stimulus location, the overall orientations and widths of these 18 tuning curves was remarkably similar. Some shifting in preferred tuning is noticed with stimulus location, but this is generally small (<50 degrees) considering the large displacements involved.

Table 1

Indexes by Experimental Class

	# Experiments	FI < 0.95	GI > 0.1	FI Average	GI Average
RD	190	158	156	0.34	1.63
ES	119	116	107	0.23	1.39
SS	184	152	143	0.37	1.33
FL	119	57	42	0.78	0.34
AP	119	36	23	0.92	0.11
NF	50	26	20	0.75	0.32

Table 2

Percentages of Significantly Different Means

	RD	ES	SS	FL	AP	NF
RD	xxx	14.3	23	34.3	16.7	36.9
ES	14.3	xxx	14	25	29.4	42.1
SS	23	14	xxx	25.9	37.5	47.1
FL	34.3	25	25.9	xxx	30	0
AP	16.7	29.4	37.5	30	xxx	0
NF	36.9	42.1	47.1	0	0	xxx

Table 3

Average Difference in Preferred Tuning (in degrees)

	RD	ES	SS	FL	AP	NF
RD	xxx	10.3	15.4	25.1	30.9	27.2
ES	10.3	xxx	10.6	20.2	21.3	25.5
SS	15.4	10.6	xxx	24.9	40	32.9
FL	25.1	20.2	24.9	xxx	31.9	16.7
AP	30.9	21.3	40	31.9	xxx	35.9
NF	27.2	25.4	32.9	16.7	35.9	xxx

Figure 1

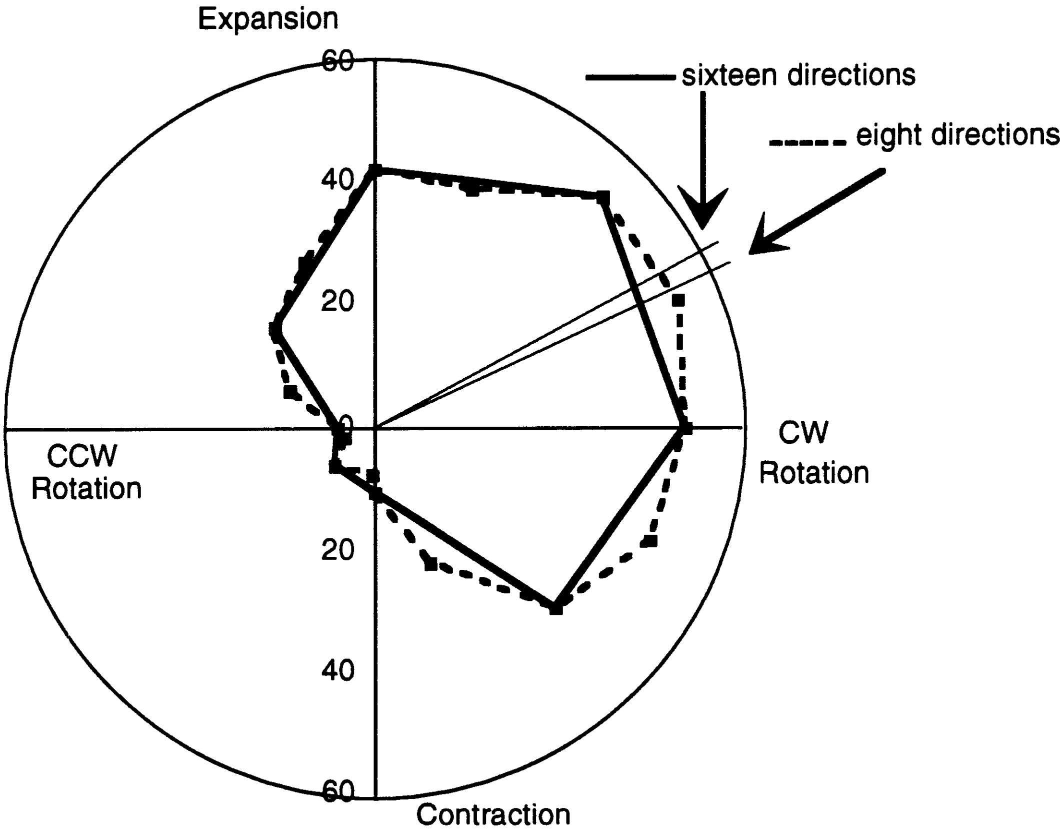

Expansion
sixteen directions
eight directions
60
40
20
0
20
40
60
CCW
Rotation
CW
Rotation
Contraction

Figure 2

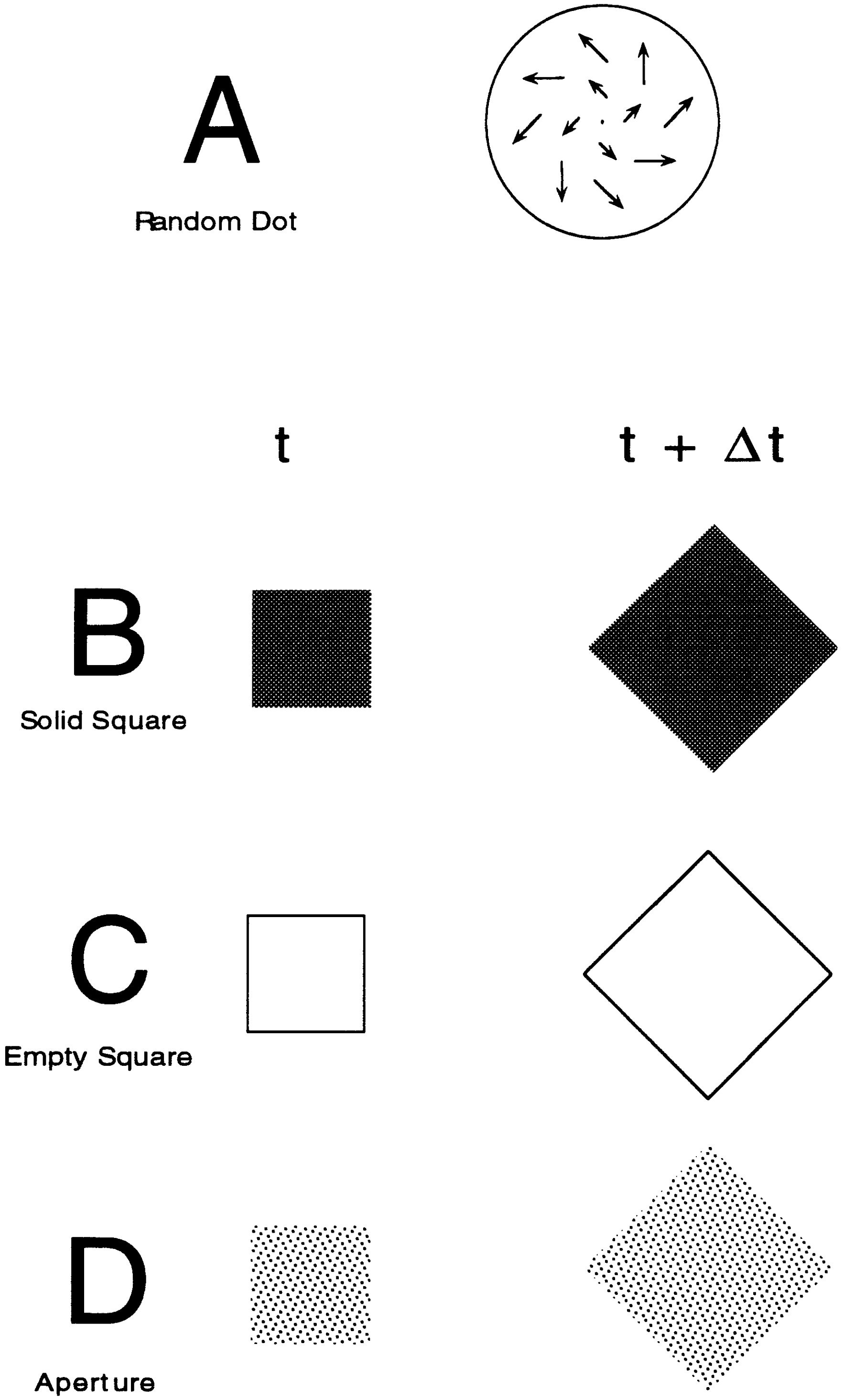
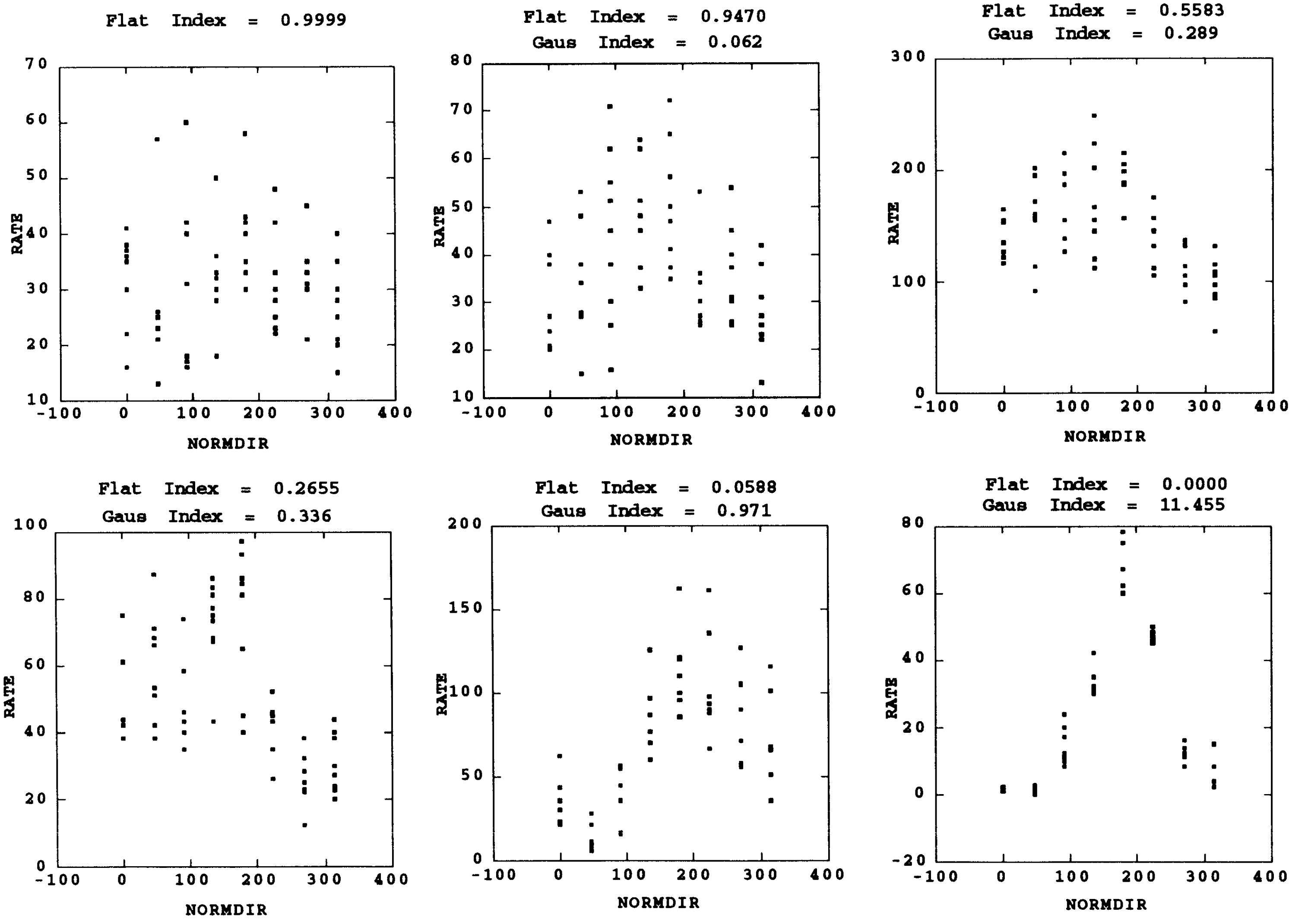


Figure 3



Flat Index = 0.9999

Flat Index = 0.9470
Gaus Index = 0.062

Flat Index = 0.5583
Gaus Index = 0.289

Flat Index = 0.2655
Gaus Index = 0.336

Flat Index = 0.0588
Gaus Index = 0.971

Flat Index = 0.0000
Gaus Index = 11.455

Figure 4

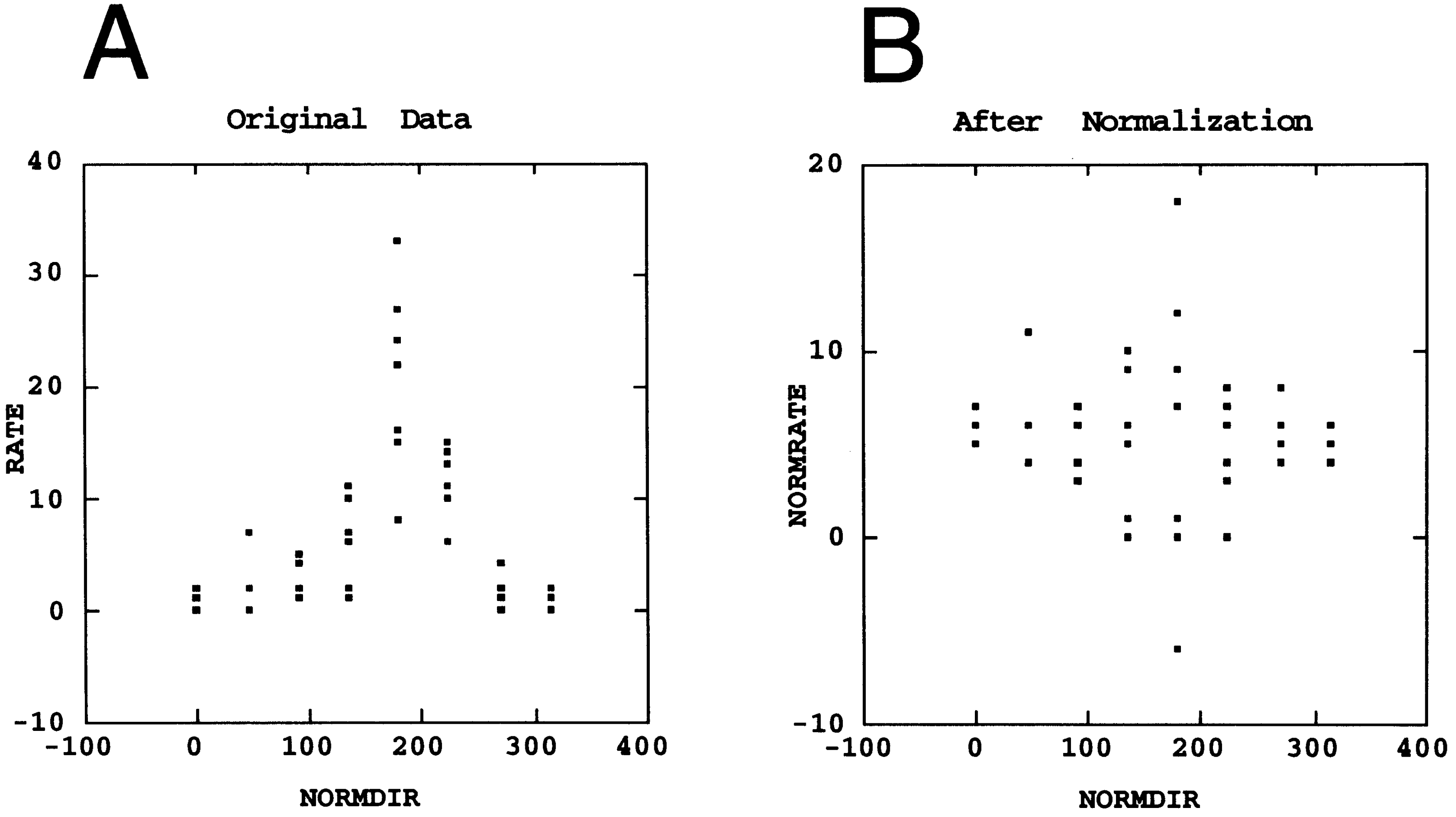
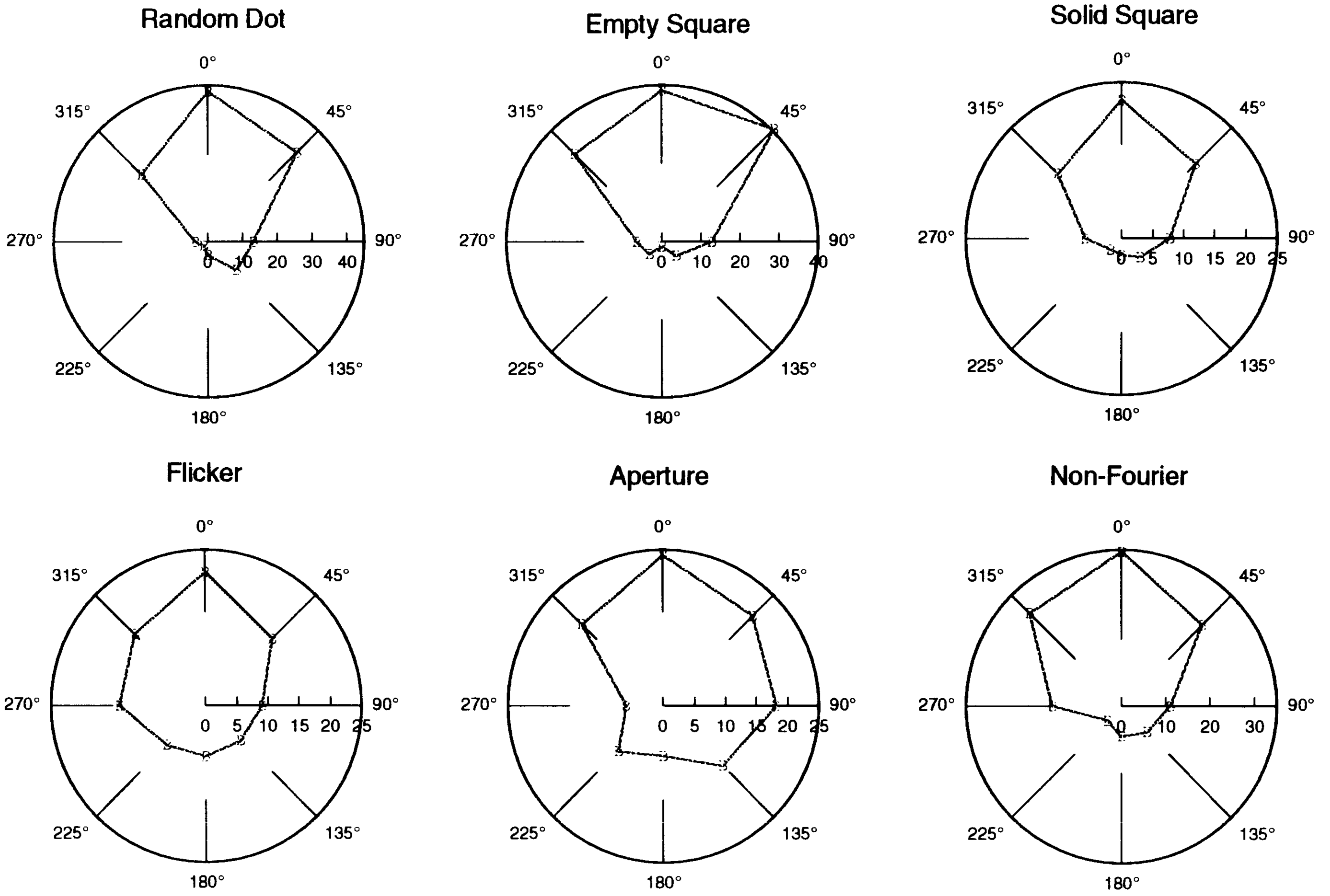

A
Original Data
RATE
40
30
20
10
0
-10
-100
0
100
200
300
400
NORMDIR
B
After Normalization
NORMRATE
20
10
0
-10
-100
0
100
200
300
400
NORMDIR

Figure 5

Unit B10800



Random Dot

Empty Square

Solid Square

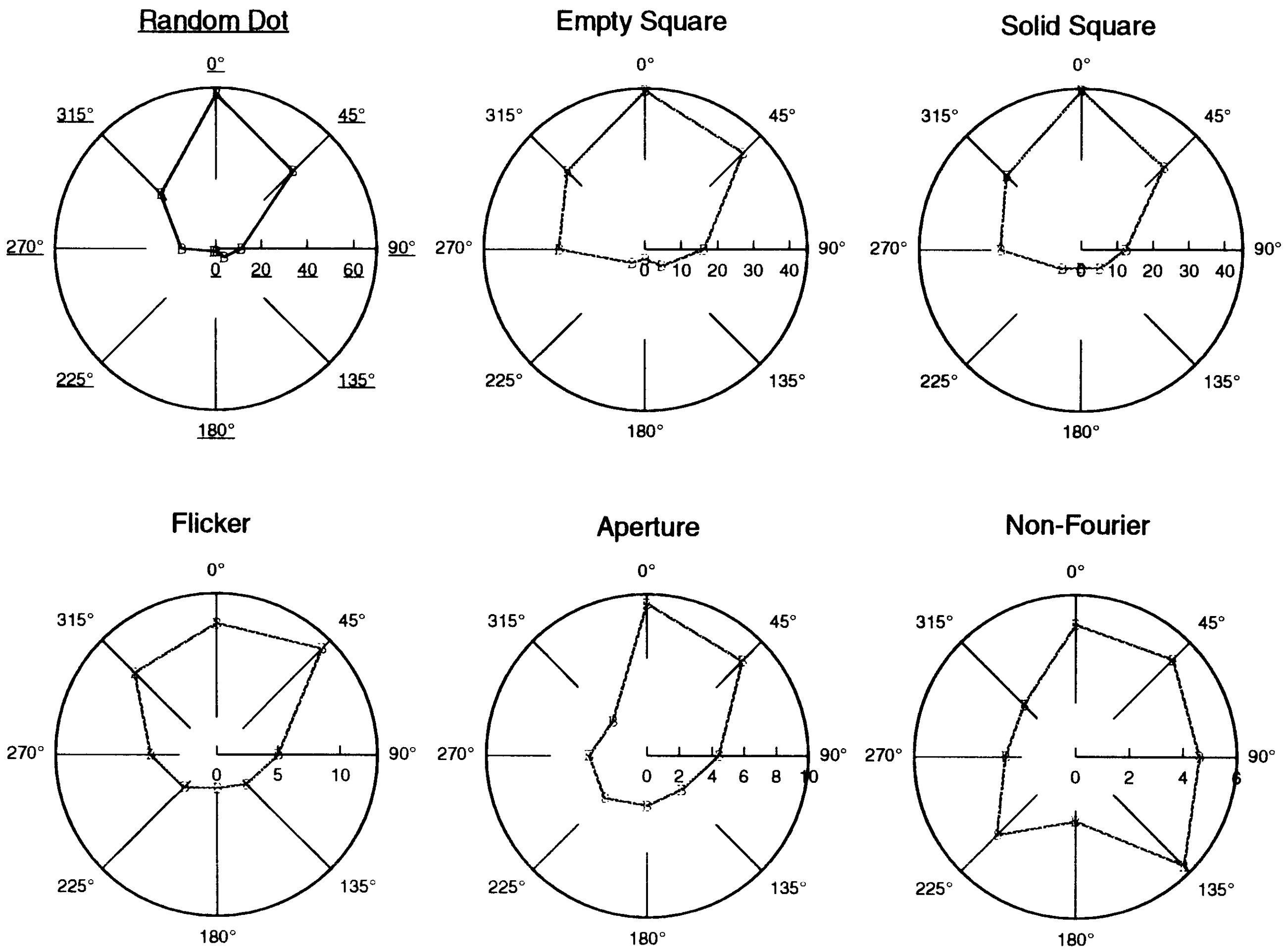
Flicker

Aperture

Non-Fourier

Figure 6

Unit B08200



Random Dot

Empty Square

Solid Square

Flicker

Aperture

Non-Fourier

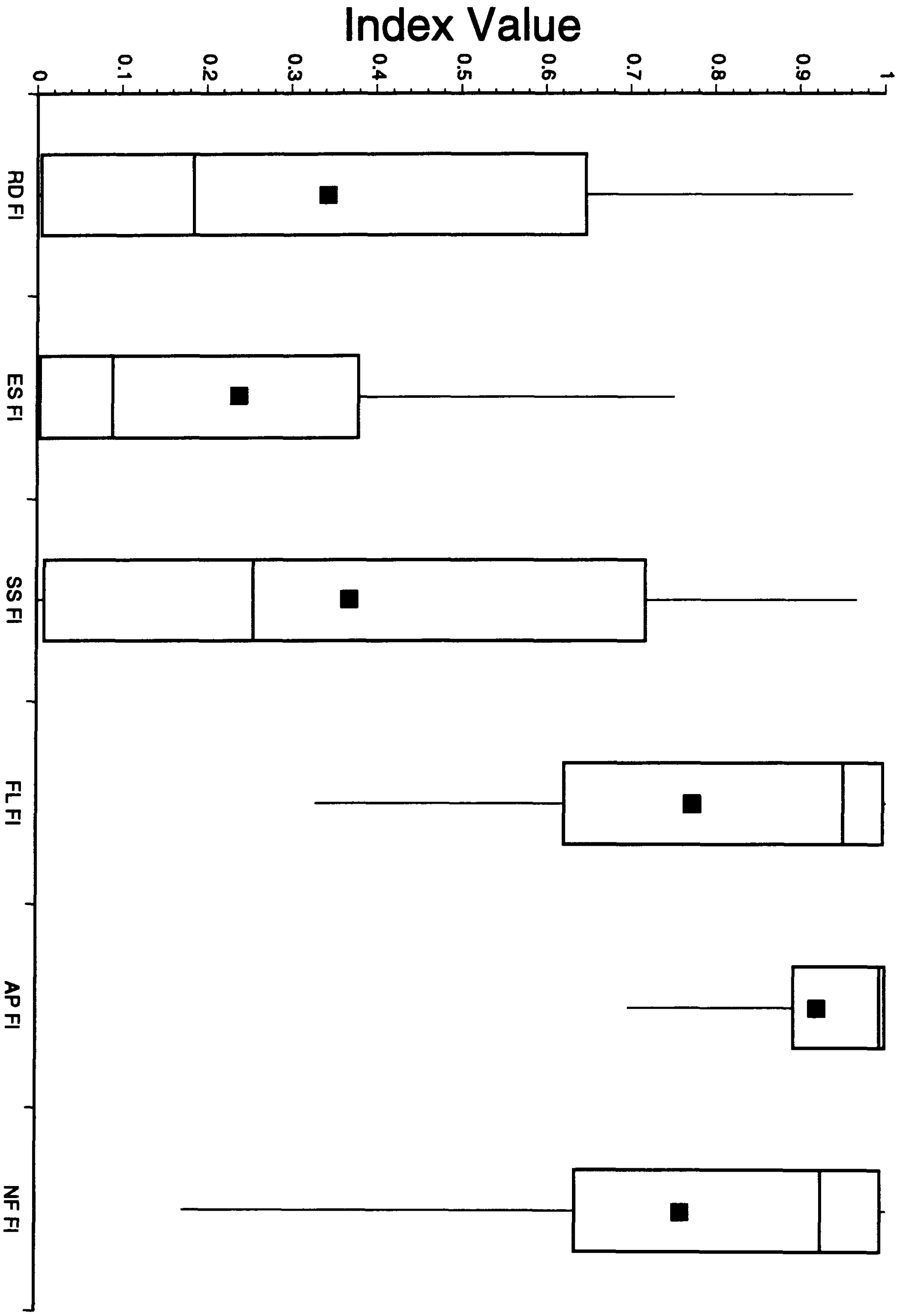

Index Value
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
RD FI
ES FI
SS FI
FL FI
AP FI
NF FI

Figure 7

Figure 8

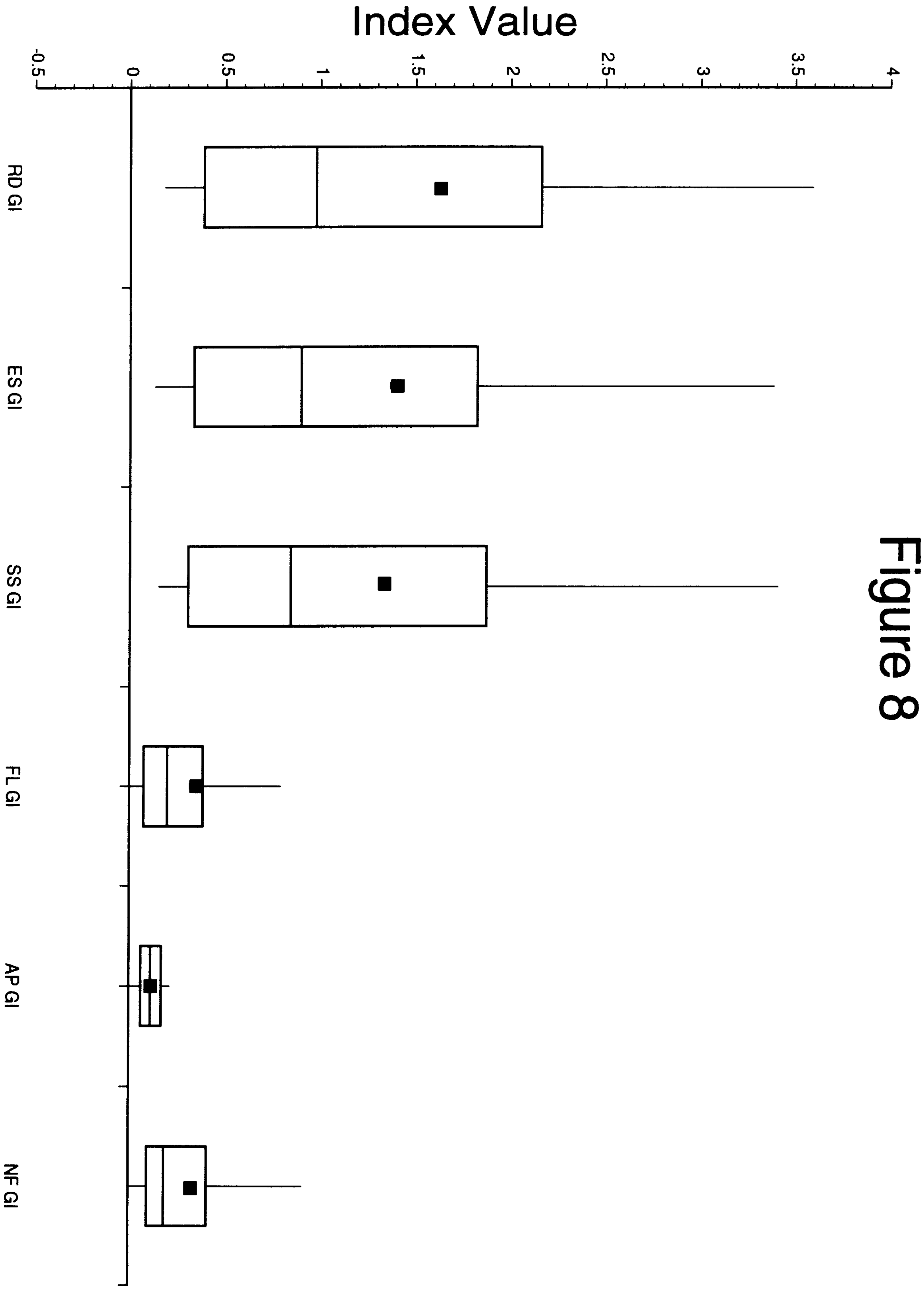


Figure 9

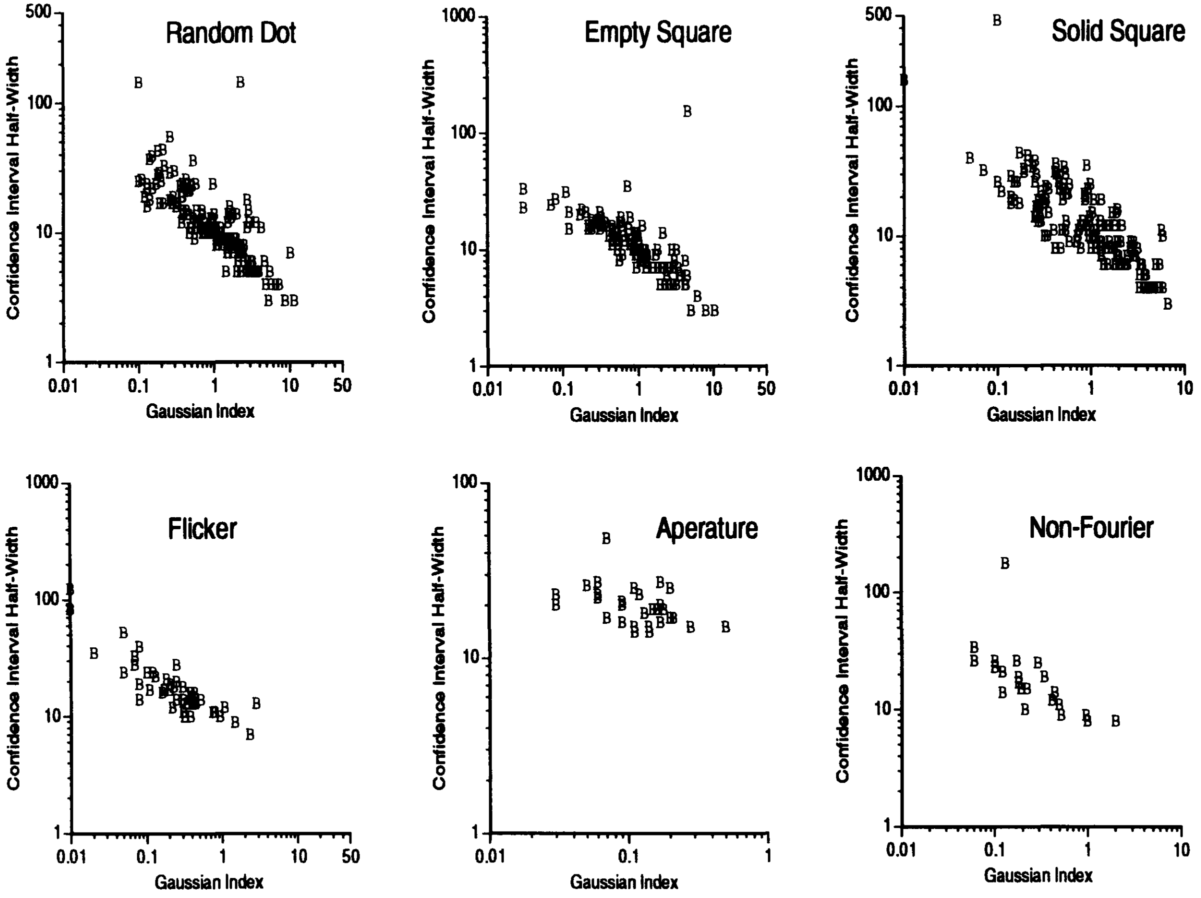
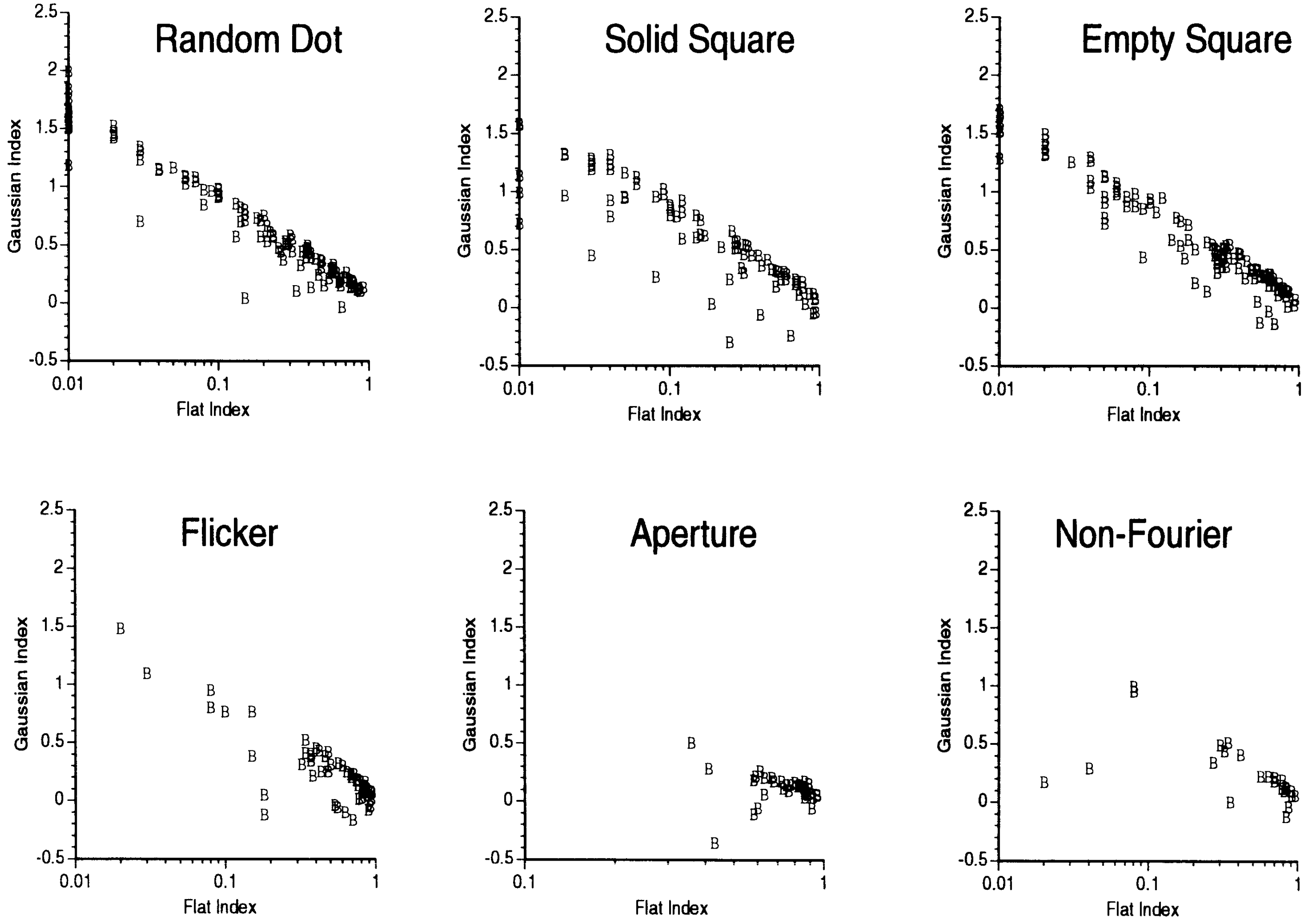

Random Dot
Empty Square
Solid Square
Flicker
Aperature
Non-Fourier
Confidence Interval Half-Width
Gaussian Index
500
100
10
1
1000
0.01
0.1
1
10
50

Figure 10


Random Dot
Solid Square
Empty Square
Flicker
Aperture
Non-Fourier
Gaussian Index
Flat Index
2.5
2
1.5
1
0.5
0
-0.5
0.01
0.1
1
B

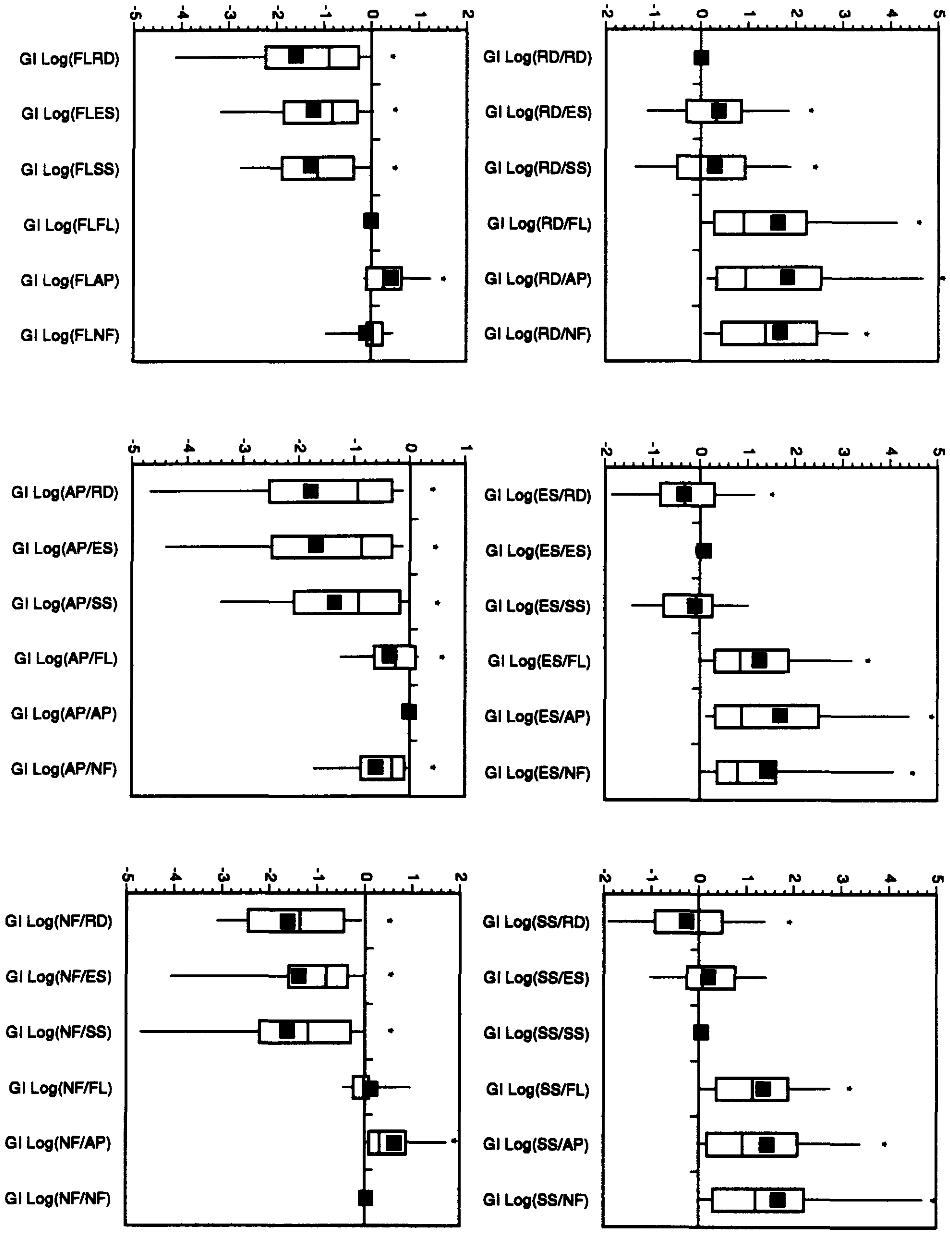
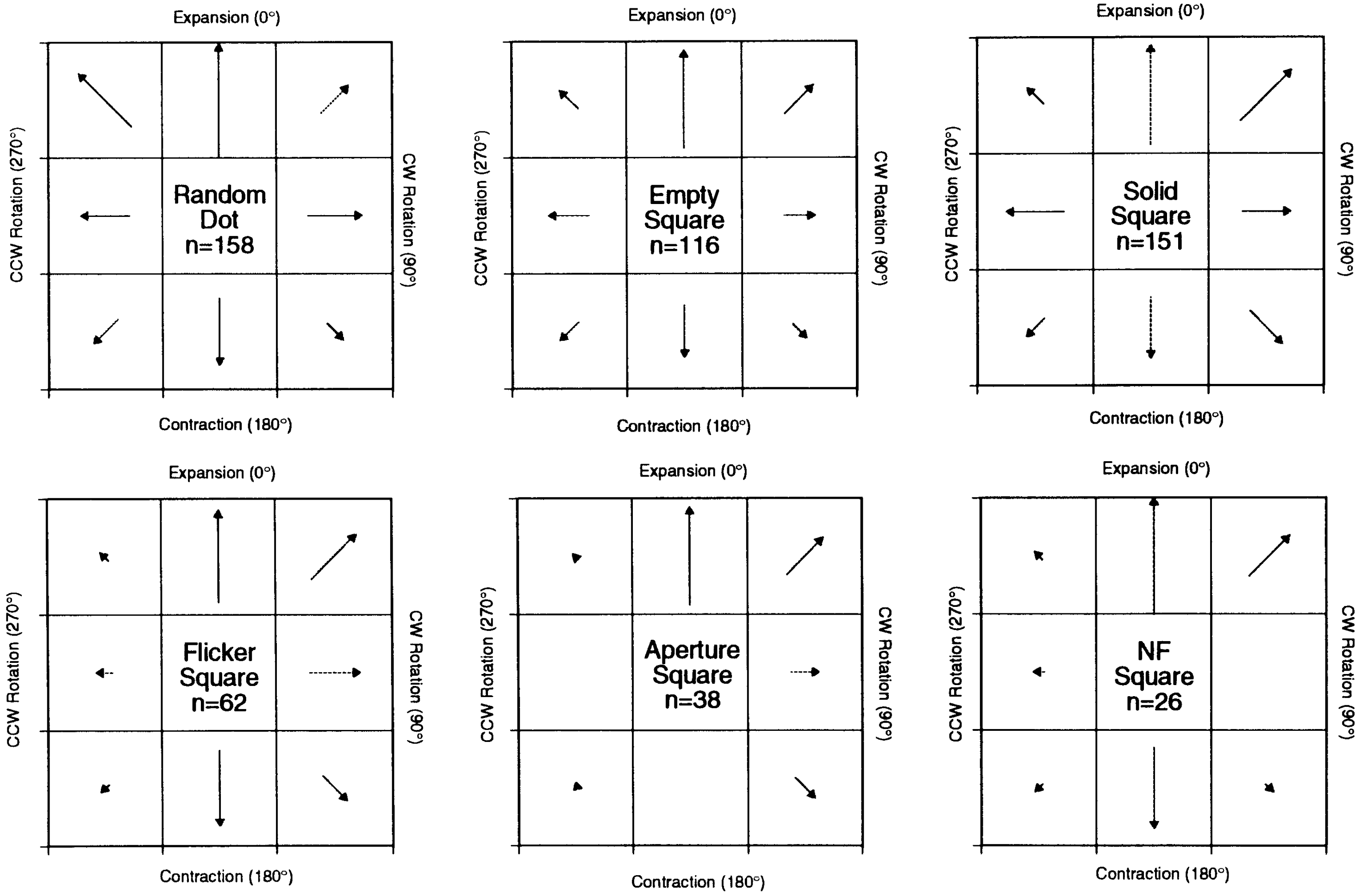

GI Log(FLRD)
GI Log(FLES)
GI Log(FLSS)
GI Log(FLFL)
GI Log(FLAP)
GI Log(FLNF)
-5
-4
-3
-2
-1
0
1
2
GI Log(RD/RD)
GI Log(RD/ES)
GI Log(RD/SS)
GI Log(RD/FL)
GI Log(RD/AP)
GI Log(RD/NF)
-2
-1
0
1
2
3
4
5
GI Log(AP/RD)
GI Log(AP/ES)
GI Log(AP/SS)
GI Log(AP/FL)
GI Log(AP/AP)
GI Log(AP/NF)
-5
-4
-3
-2
-1
0
1
GI Log(ES/RD)
GI Log(ES/ES)
GI Log(ES/SS)
GI Log(ES/FL)
GI Log(ES/AP)
GI Log(ES/NF)
-2
-1
0
1
2
3
4
5
GI Log(NF/RD)
GI Log(NF/ES)
GI Log(NF/SS)
GI Log(NF/FL)
GI Log(NF/AP)
GI Log(NF/NF)
-5
-4
-3
-2
-1
0
1
2
GI Log(SS/RD)
GI Log(SS/ES)
GI Log(SS/SS)
GI Log(SS/FL)
GI Log(SS/AP)
GI Log(SS/NF)
-2
-1
0
1
2
3
4
5

Figure 11

Figure 12


Expansion (0°)
CCW Rotation (270°)
CW Rotation (90°)
Contraction (180°)
Random
Dot
n=158
Expansion (0°)
CCW Rotation (270°)
CW Rotation (90°)
Contraction (180°)
Empty
Square
n=116
Expansion (0°)
CCW Rotation (270°)
CW Rotation (90°)
Contraction (180°)
Solid
Square
n=151
Expansion (0°)
CCW Rotation (270°)
CW Rotation (90°)
Contraction (180°)
Flicker
Square
n=62
Expansion (0°)
CCW Rotation (270°)
CW Rotation (90°)
Contraction (180°)
Aperture
Square
n=38
Expansion (0°)
CCW Rotation (270°)
CW Rotation (90°)
Contraction (180°)
NF
Square
n=26

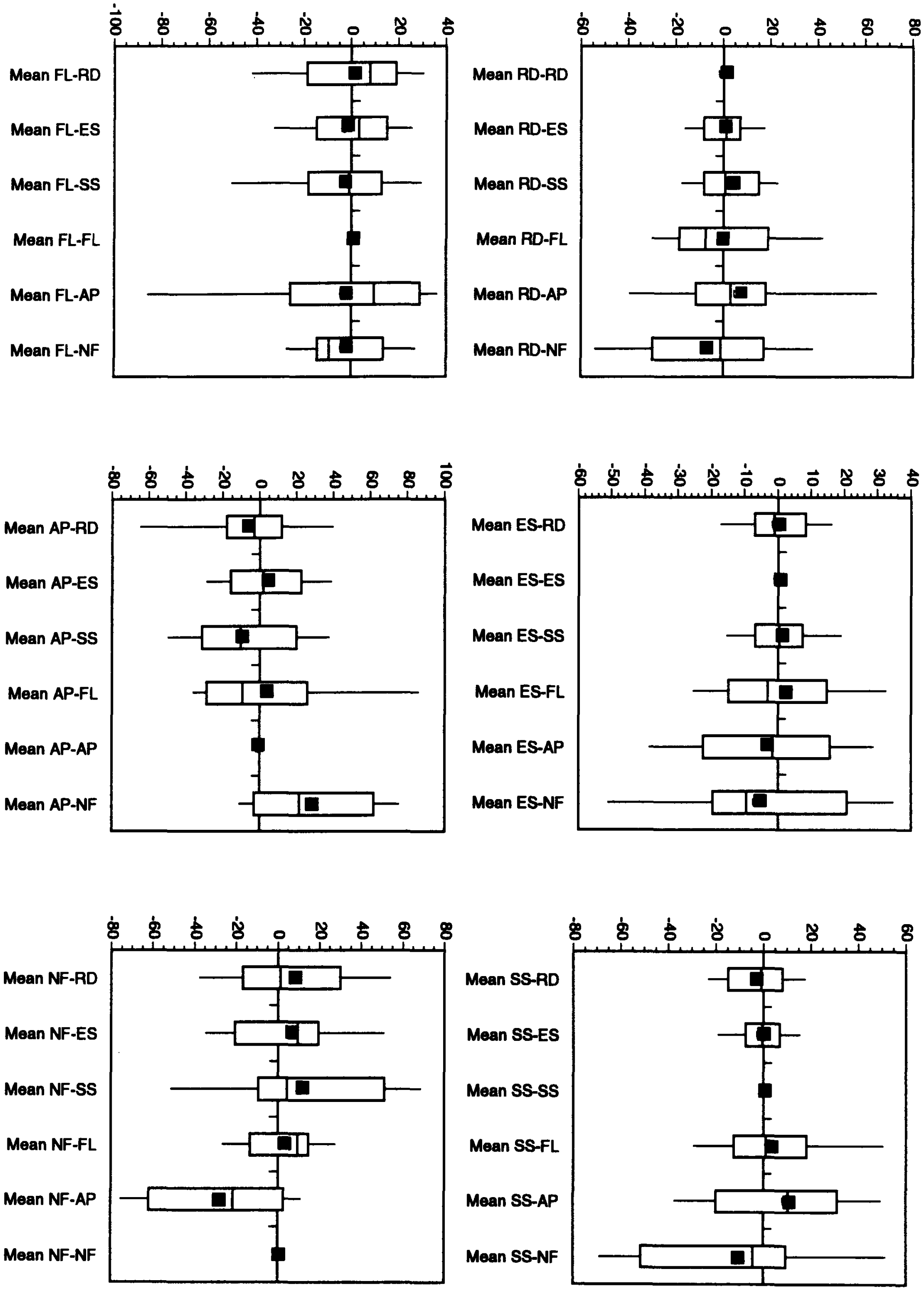

-100
-80
-60
-40
-20
0
20
40
Mean FL-RD
Mean FL-ES
Mean FL-SS
Mean FL-FL
Mean FL-AP
Mean FL-NF
-60
-40
-20
0
20
40
60
80
Mean RD-RD
Mean RD-ES
Mean RD-SS
Mean RD-FL
Mean RD-AP
Mean RD-NF
-80
-60
-40
-20
0
20
40
60
80
100
Mean AP-RD
Mean AP-ES
Mean AP-SS
Mean AP-FL
Mean AP-AP
Mean AP-NF
-60
-50
-40
-30
-20
-10
0
10
20
30
40
Mean ES-RD
Mean ES-ES
Mean ES-SS
Mean ES-FL
Mean ES-AP
Mean ES-NF
-80
-60
-40
-20
0
20
40
60
80
Mean NF-RD
Mean NF-ES
Mean NF-SS
Mean NF-FL
Mean NF-AP
Mean NF-NF
-80
-60
-40
-20
0
20
40
60
Mean SS-RD
Mean SS-ES
Mean SS-SS
Mean SS-FL
Mean SS-AP
Mean SS-NF

Figure 13

Figure 14

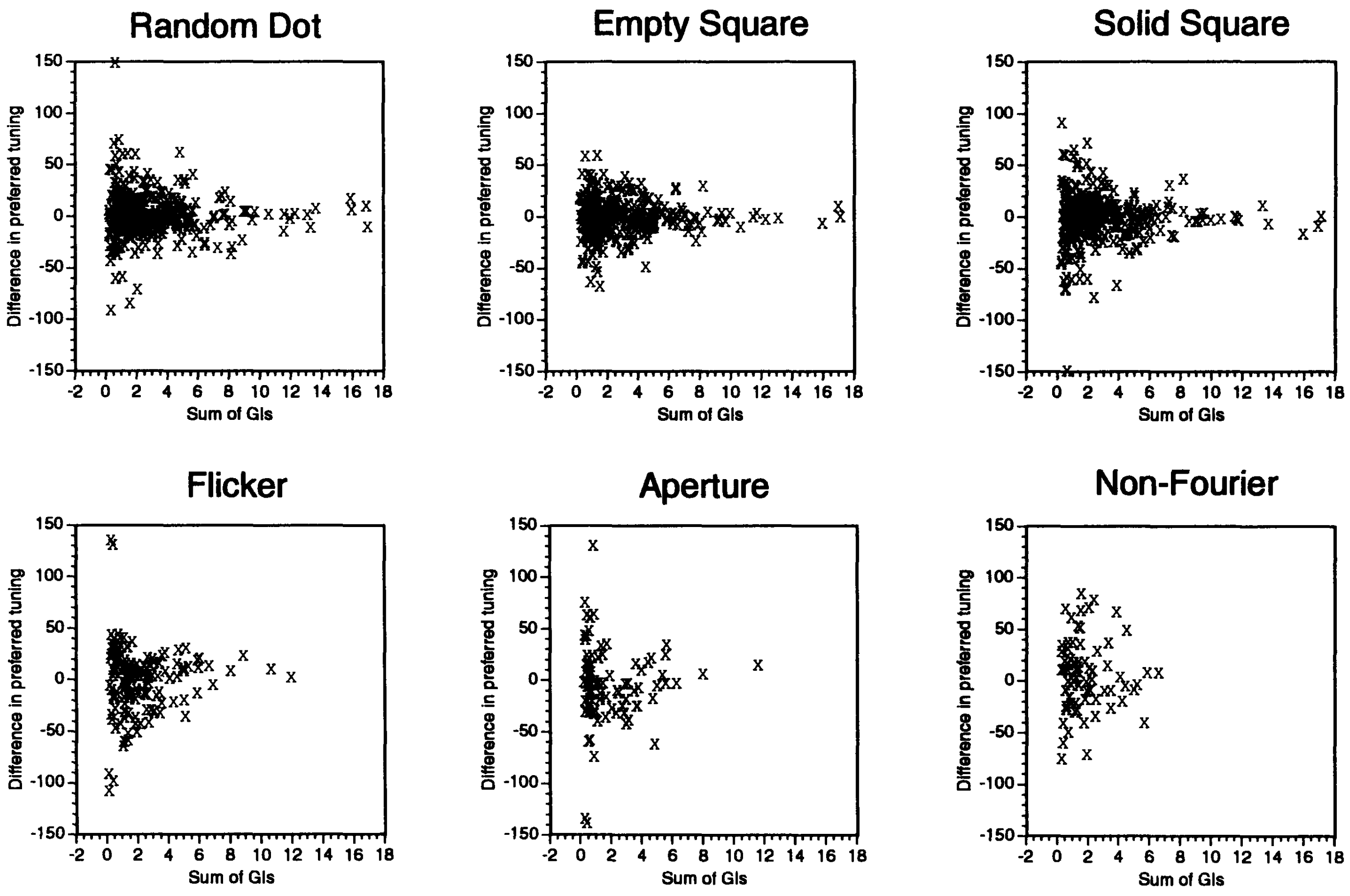

Random Dot
Empty Square
Solid Square
Flicker
Aperture
Non-Fourier
Difference in preferred tuning
150
100
50
0
-50
-100
-150
Sum of GIs
-2
0
2
4
6
8
10
12
14
16
18

Figure 15

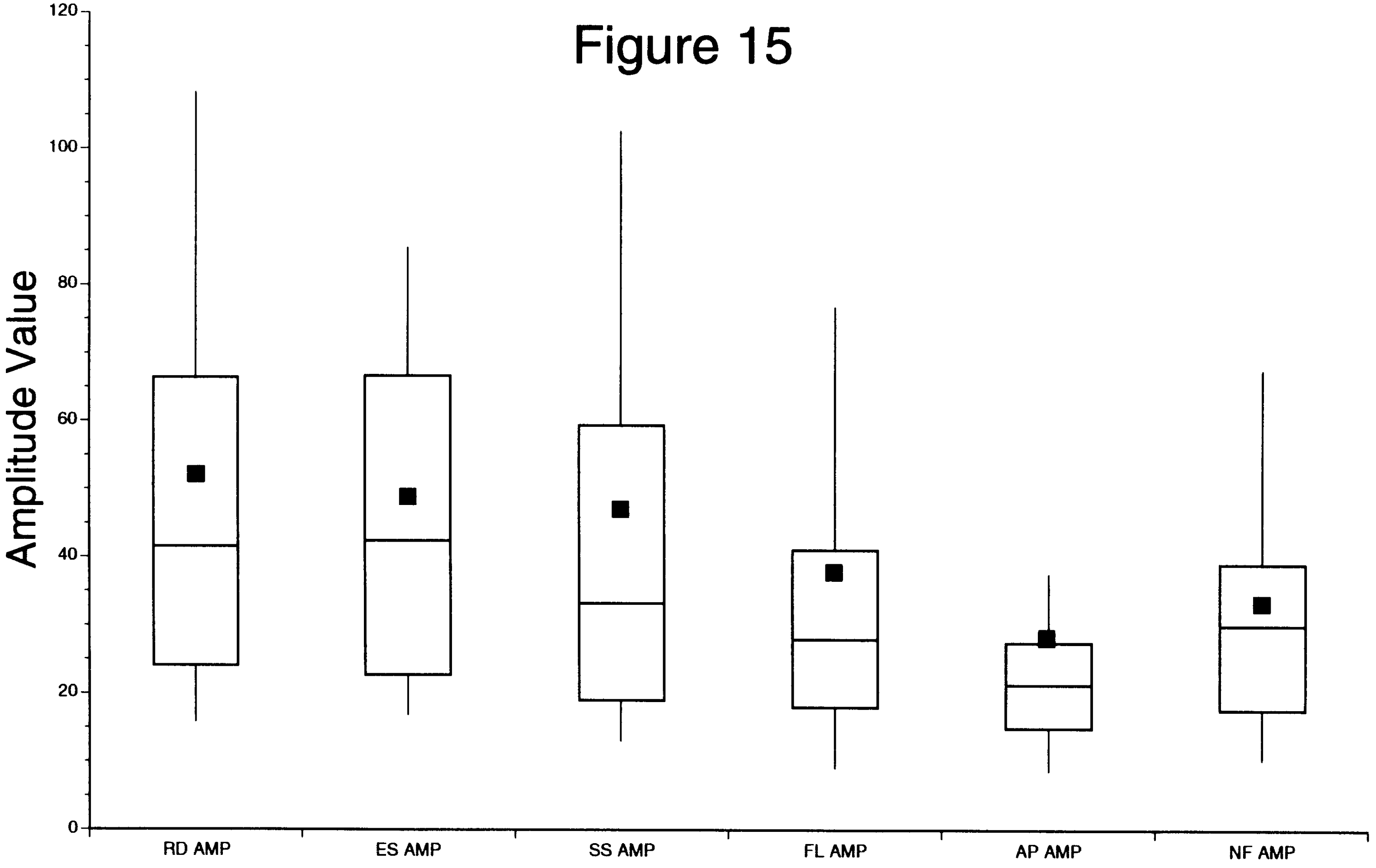
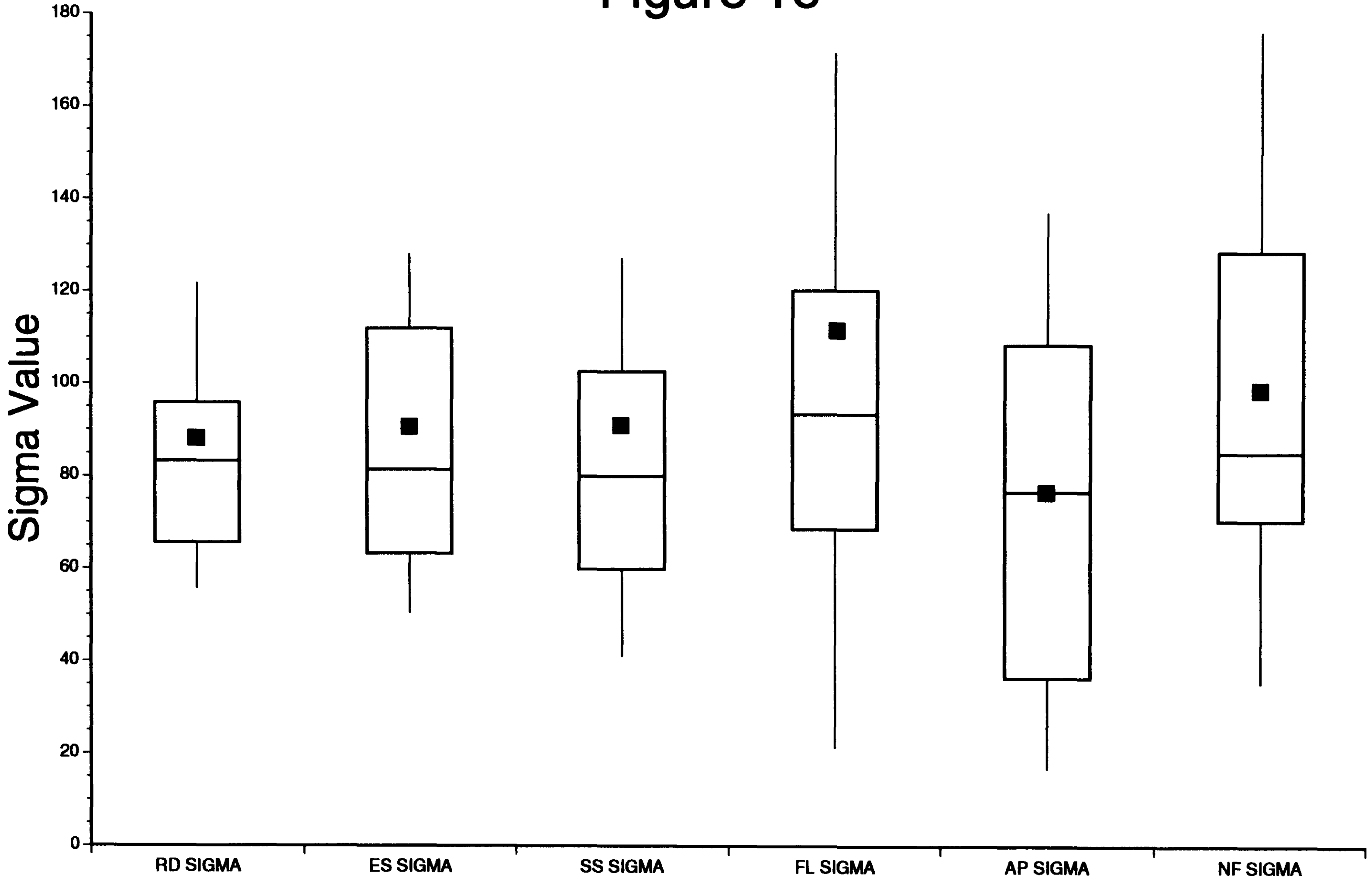

120
100
80
60
40
20
0
Amplitude Value
RD AMP
ES AMP
SS AMP
FL AMP
AP AMP
NF AMP

Figure 16


Sigma Value
180
160
140
120
100
80
60
40
20
0
RD SIGMA
ES SIGMA
SS SIGMA
FL SIGMA
AP SIGMA
NF SIGMA

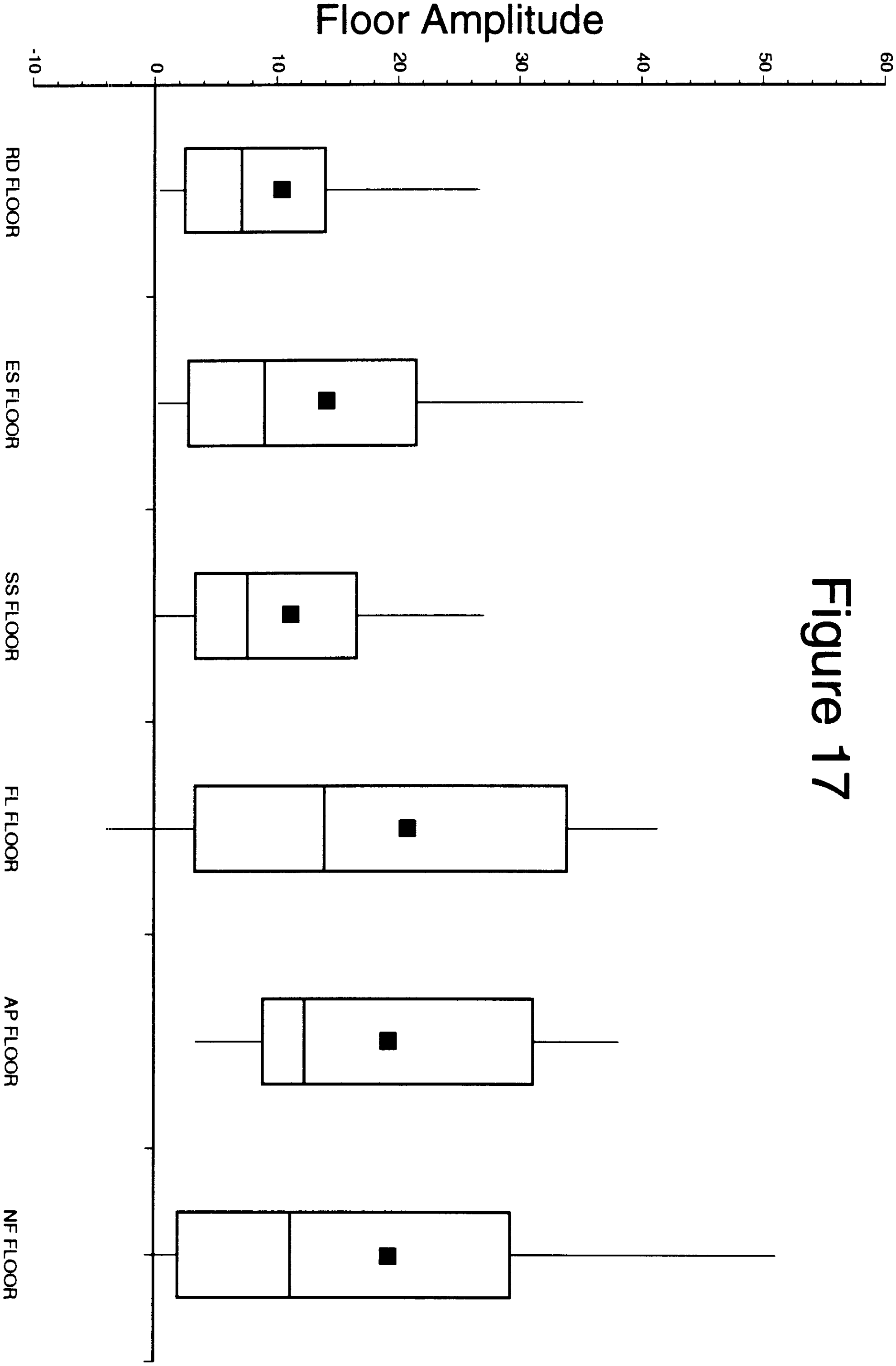

Floor Amplitude
-10
0
10
20
30
40
50
60
RD FLOOR
ES FLOOR
SS FLOOR
FL FLOOR
AP FLOOR
NF FLOOR

Figure 17

Figure 18

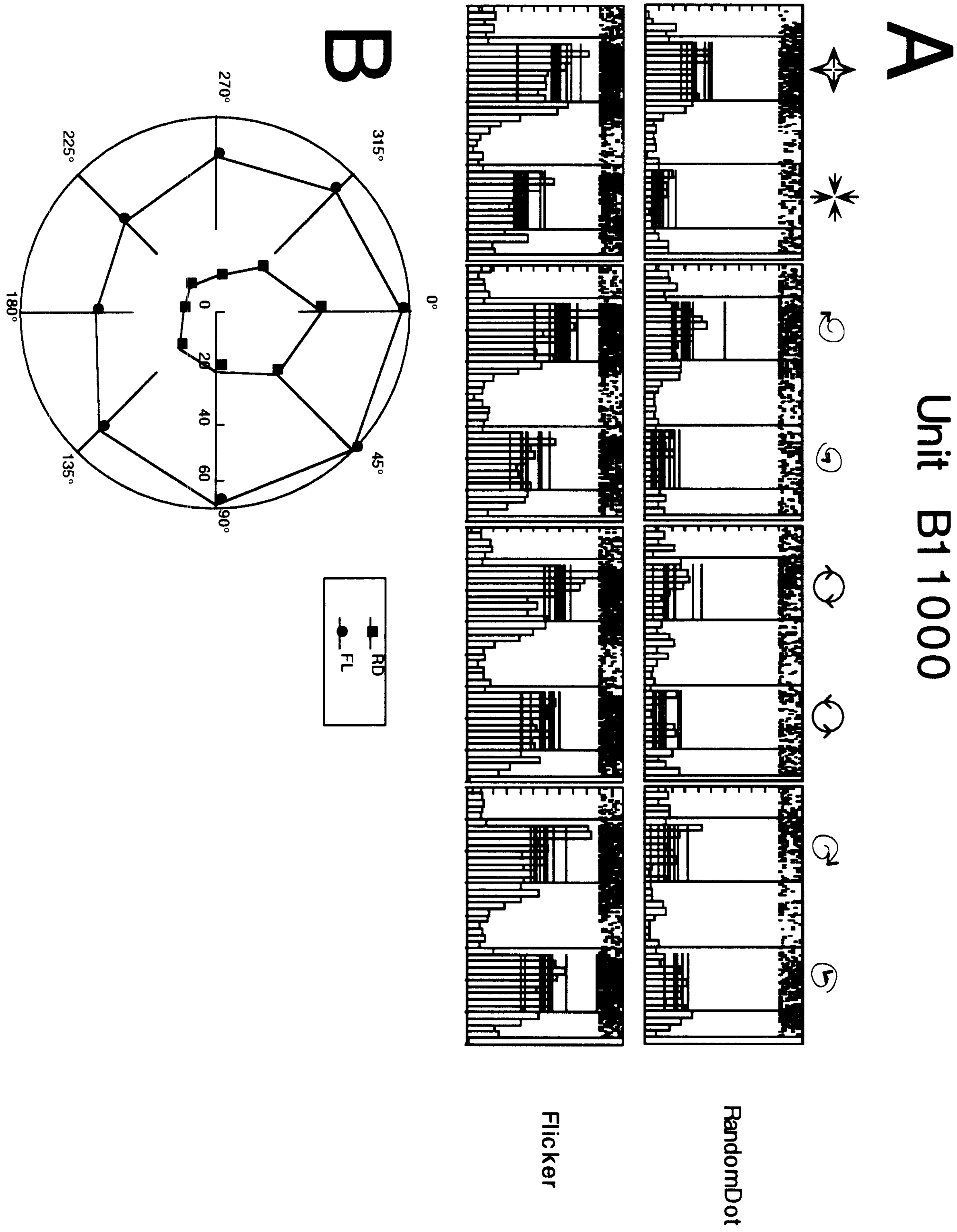

A
Unit B11000
RandomDot
Flicker
B
270°
315°
0°
45°
90°
135°
180°
225°
0
20
40
60
RD
FL

Figure 19

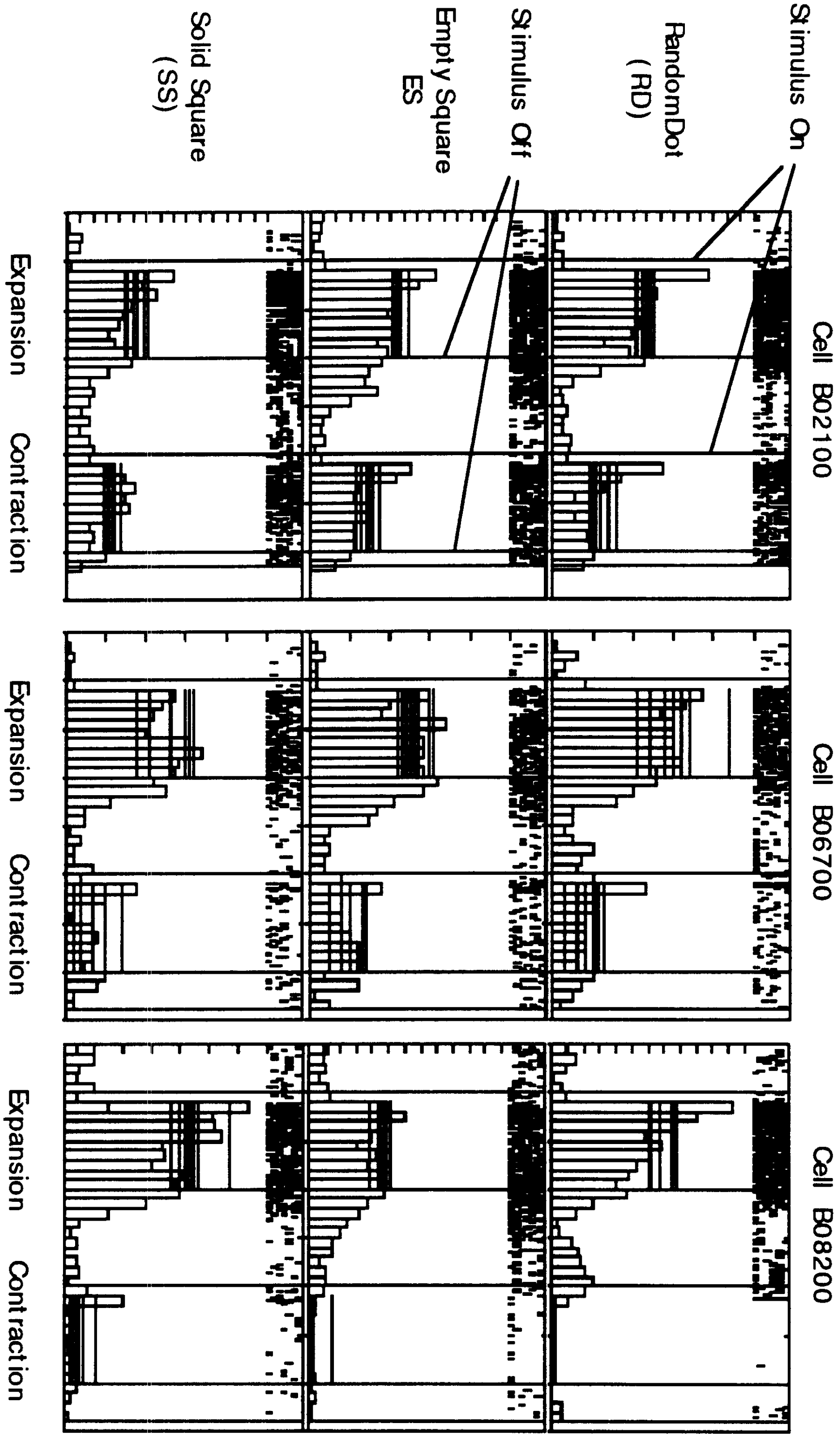
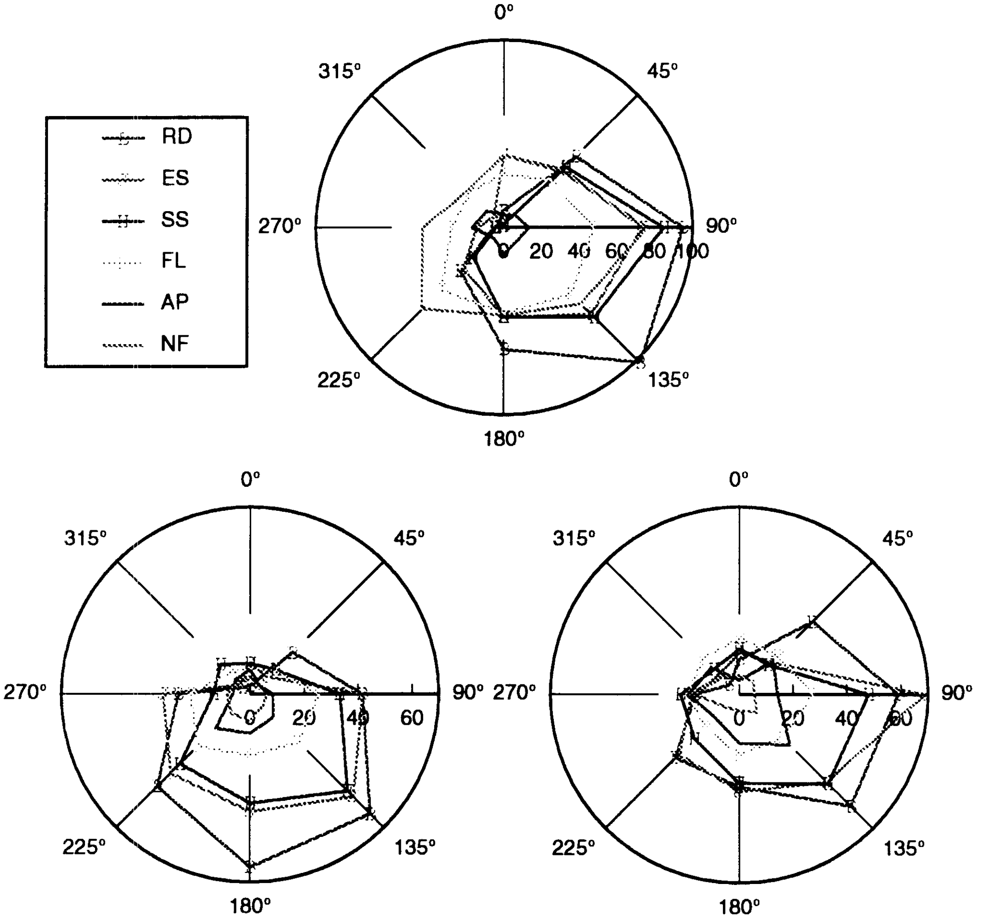

Stimulus On
RandomDot
(RD)
Stimulus Off
Empty Square
ES
Solid Square
(SS)
Cell B02100
Cell B06700
Cell B08200
Expansion
Contraction
Expansion
Contraction
Expansion
Contraction

Figure 20

Unit B11300



Chapter 3

Functional Organization in Macaque Superior Temporal Sulcus With Regards to Complex Motion Patterns

ABSTRACT

The superior temporal sulcus (STS) of the macaque monkey contains multiple visual areas (Ungerleider & Desimone, 1986). Neurons within these regions have been shown to respond selectively to wide-field, complex motion patterns such as expansion, contraction, and rotation. Units driven by these stimuli are believed to be located within cortical areas MST. Single-unit recording studies in this region have provided evidence for a systematic spatial organization with respect to neuronal specificity. Cells with similar tuning properties were found clustered together into columns or slabs extending through multiple cortical layers. In this study, we use a double-label 2-deoxyglucose technique in an attempt to directly visualize this functional organization. This investigation is unique in that it is the first time the double-label method has been attempted in the awake behaving animal. Two monkeys were used, both of which viewed expansion during one labeling period and either contraction or rotation in the other. Two different functional arrangements of neurons have been proposed for this region. The stimulus pairs chosen were designed to help decide between these two proposals. The pattern of activity visible on the autoradiograms suggests a periodic arrangement of columns with alternating stimulus selectivity, with units tuned for expansion and contraction being maximally separated within this mosaic. Although the region of interdigitating stimulus specificity was clearly visible on the floor and posterior bank of the STS, single unit mapping, cytochrome staining, and myelin staining we unable to unambiguously assign borders to the various cortical regions within the STS. Despite this, we believe the evidence firmly supports partial overlap of MT with the region of differential labeling. However, although there is evidence to support MST overlapping the residual portion of the area, this remains to be firmly established.

INTRODUCTION

Considerable evidence has accumulated supporting the modular organization of mammalian cortex. In the 1950s, Mountcastle, working in cat somatosensory cortex, reported that units specific for different sensory sub-modalities such as light touch and deep touch were clustered according to response selectivity (Mountcastle, 1957). Neurons with similar response profiles were encountered while advancing a recording electrode perpendicular to the cortical surface, but when the electrode was advanced tangentially, specificity gradually shifted. This organization appeared to be columnar, with the axis of the iso-tuning columns oriented perpendicular to the surface of the cortex and extending through most cortical layers.

The tendency for neurons of like specificity to cluster has been observed throughout cortex. In area V1 (cat areas 17, 18) a columnar organization appears to be present with regards to orientation selectivity, ocular dominance, and direction selectivity (Hubel & Wiesel, 1962; Blasdel, 1992; Berman et al, 1987; Bartfeld et al, 1992; Redies et al, 1990). Furthermore, functional NMR, optical imaging, single unit recording, and 2-deoxyglucose studies have all reported a similar architecture, strongly arguing against the possibility that these effects are artifacts associated with a particular technique. How the functional maps for these three stimulus dimensions are organized relative to one is only beginning to be understood (Blasdel, 1992). Within a single stimulus domain, it is clear that response specificity changes in a gradual and systematic way moving across cortical columns. Segregation of neurons by preferred tuning along a single stimulus dimension has also been observed for auditory frequency in the inferior colliculus (Serviere et al, 1984; O'Neill et al, 1989) and odor in olfactory cortex (Skeen, 1977), to name just two examples. Regardless of the cortical region considered or the particular technique employed, functional columns of 0.2 - 0.5 mm form a mosaic that extends through most layers of the cortex, with a tendency to see coarser columnar organization downstream from primary sensory cortex. It is likely that this organization plays an important role in cortical function.

In this study, we examined the columnar organization of cortical regions in the superior temporal sulcus (STS) thought to be involved in processing visual motion information (for a review, see Albright, 1993). In primates, analysis of visual motion first occurs in area V1, where directional tuning has been found in a large number of units. Output from these neurons feeds into what has been described as the "where" cortical processing stream, that is thought to be involved in analyzing positional and motion cues (DeYoe et al, 1988). The STS contains several anatomic regions which are thought to be intermediate processing stations in this system. The most studied of these regions is area MT, which receives a direct projection from V1, has small receptive fields, and contains neurons which display directional selectivity (Maunsell et al, 1983; Albright, 1984; Raiguel et al, 1993). A columnar arrangement with regards to directional selectivity has been proposed for both primary visual cortex (Berman et al, 1987) and area MT (Albright, 1984).

The area receiving the strongest projection from area MT is the adjacent cortical region, MST. While MT occupies part of the anterior bank and fundus of the STS, MST has been identified just anterior, on the floor and upper bank of the sulcus (Desimone et al, 1986; Ungerleider et al, 1986; Nawrot et al, 1989). MST itself has been divided, mostly on physiological criteria, into at least two subdivision, a lateral (MSTl) and a dorsal (MSTd) area. MSTl typically has neurons with eccentrically placed receptive fields which give directionally tuned responses to translational motion stimuli (Tanaka et al, 1993; Komatsu et al, 1988; Wurtz et al, 1988). The size of these units' receptive fields is approximately that of an MT cell at an equivalent receptive field eccentricity. Area MSTd, in contrast, consists of units with large (>20 degree) diameter receptive fields that generally include the fovea and often extend into the ipsilateral receptive field (Saito et al, 1986; Graziano et al, 1994; Tanaka et al, 1989ab; Duffy et al, 1991ab). Like cells in areas MSTl and MT, the majority of neurons in MSTd are directionally tuned for straight (i.e. translational) motion. However, many units in MSTd are also tuned for more complex motion patterns such as expansion, rotation, and contraction. In addition, unlike some units in MSTl and MT, neurons in MSTd do not have inhibitory surrounds and generally respond best to wide-field "optical flow" motion stimuli. Making the physiology even more complicated, it is common for neurons in MSTd to be tuned simultaneously for both translational and complex motion. For example, units have been found that prefer both upward motion to downward motion and expansion to contraction (Graziano et al, 1994).

The types of stimuli most effective in driving MSTd units are similar to the motion patterns produced on the retina as the result of observer movement through the environment. Since first proposed by Gibson, it has been recognized that this "optical flow" contains the necessary information for recovering direction of heading. It has also been proposed that analysis of these complex motion patterns is important for the representation of object motion in the environment (Graziano, et al 1994). For example, units tuned for expansion may be important in representing the flight of an approaching object or the pattern of ripples produced after dropping a stone in water. Furthermore, much psychophysical evidence exists suggesting that we extract expansion and contraction in the visual scene at a very fundamental level (Regan, 1986; Freeman et al, 1992). Given the special status afforded to the analysis of complex motion patterns, we were interested in determining how this stimulus attribute was represented in the functional anatomy of MSTd. Previous single-unit studies have reported "clumping" of neurons in this region according to response selectivity (Duffy et al, 1991a). One group has shown physiological evidence for units tuned for expansion, contraction, and the two directions of rotation segregated into separate cortical columns (Lagae et al, 1994). In their model, columns for expansion and contraction are maximally separated, as are columns for clockwise and counter-clockwise rotation.

We decided to test this hypothesis directly by using the double-labeled 2-deoxyglucose method (2-DG). The deoxyglucose technique relies on the assumption that there exists a positive correlation between the average activity of neurons in a localized region of tissue and the rate of glucose utilization (Theurich et al, 1984; Sokoloff et al, 1977; Friedman et al, 1987; Sokoloff, 1977). After injecting a radioactive analog of glucose (14C-deoxyglucose) into the organism at the start of a stimulus period, neurons involved in maintaining the neural state induced by the stimulus take up more label than the same neurons without the stimulus present. For example, cells in V1 orientation columns tuned to vertical contours transport more labeled glucose when vertical compared to horizontal lines are displayed. This differential glucose uptake can be visualized by slicing the tissue into thin sections and then exposing the sections to film sensitive to the appropriate isotope. Iso-orientation columns in V1 have been visualized in this way as a periodic mosaic of differentially labeled columns or slabs (LeVay et al, 1975; Redies et al, 1990; Singer, 1981; Albus et al, 1984; Schoppmann et al, 1981; Albus, 1979; Humphrey et al, 1980).

The standard 2-DG technique described above has several inherent limitations. For example, the pattern of differential label uptake recorded in the autoradiograms may not reflect the effect of the stimulus presented to the animal but instead represent intrinsic background variability in the glucose utilization across local regions of cortex. The cytochrome oxidase blobs of primary visual cortex are a classic example of this — even after enucleation, 2-DG label will preferentially be taken up in these metabolically more active regions (Wong-Riley, 1979). Although this problem can be partially solved by using both a control and an experimental group, another problem remains. It is difficult to determine which aspect/s of the stimulus are giving rise to the pattern of labeling reported in these studies. For the case of V1 orientation columns described above, the columnar labeling pattern may unintentionally reflect the anatomical organization of units specific for the color, spatial frequency content, or size of the inducing stimulus.

The double-label 2-DG technique offers an elegant solution to these problems. By using two different metabolic isotopes whose labeling patterns can be recovered independently, two different stimulus conditions can be compared within the same organism, allowing each specimen to serve as its own control (Friedman et al, 1989; Redies et al, 1987). If only one stimulus dimension (e.g. color, direction of motion, spatial frequency, etc.) is different for the two stimulus periods, differences in the labeling patterns must be a consequence of changing the single stimulus aspect manipulated. If the region of the brain examined has a columnar organization with respect to the stimulus dimension varied between exposures, these differences appear as areas of interdigitation in images where the labeling patterns of both exposures are simultaneously visible (e.g. see difference images in METHODS). Any correlated pattern of labeling in the two images reflects either intrinsic base-line variability in metabolic activity or differential specificity for a stimulus dimension along which the two stimulus patterns are identical. Therefore, we predict that if regions within the STS have a functional organization with respect to complex motion patterns, these areas should appear as regions of interdigitation when the monkey views the appropriate motion patterns during the two labeling periods.

Two macaque monkeys were used for this experiment. In one animal (macaque 92-1), the two labeling periods compared counter-clockwise rotation and expansion. For the other animal (macaque 92-2), expansion and contraction patterns served as the stimuli. As discussed above, one study (Lagae et al, 1994) suggested that expansion specific and contraction specific cell clusters are maximally separated, with rotation tuned units interspersed between these two types of columns. Comparing motion patterns which activate units that are maximally spatially separated should produce labeling patterns which are 180 degrees out of phase in the two exposures. If their model is correct, the two labeling patterns should show a greater correlation for expansion/rotation than for expansion/contraction pairing. A unit without specificity for complex motion but tuned to translational (linear) motion would be expected to give strong responses to our stimuli if the receptive fields of these neurons overlapped an approximately linear portion of our patterns. This would occur for cells with fairly small receptive fields viewing a peripheral part of the stimulus. A substantial portion of area MT meets these criteria and we therefore expected differential labeling in this region as well as in MST.

METHODS

Stimuli

The motion patterns used in this study were essentially identical to those developed by Graziano, et al (1994), except that our stimuli were much larger (40 degrees). Briefly, these stimuli are limited life-time random dot motion patterns (Figure 1), displayed at a refresh rate of 60 Hz. Every screen refresh cycle, the positions of the dots on the display are shifted, consistent with the global motion being simulated. Each frame contains 200 dots within the circular window of the display. Each dot remained on the screen for 333 msec (20 frames) before being randomly reassigned a new position within the display window and a new trajectory consistent with its new location. The dots were relocated asynchronously, to avoid a coherent flickering of the stimulus. If the dot moved outside the bounds of the display window, it was immediately assigned a new, random location within the display and a new trajectory. The speed of each dot was a linear function of its distance from the center of the display, given by the formula speed = 0.2 * (distance from the center of the stimulus), the units being arbitrary. During each dot life-time the velocity of the dot was held constant, eliminating any element of acceleration or curvature from its path. The direction of motion for each dot was determined by the type of global motion desired; i.e. expansion requires each dot to be moving directly away from the center of the stimulus; rotation results from rotating each of the expansion component motion vectors 90 degrees and contraction from rotating the dots another 90 degrees in the same direction. Therefore, one motion type could be transformed into another without any change in the overall velocity distribution — only the global organization of these local motion components changes. Graziano, et al (1994) demonstrated that patterns designed in this way produce strong, tuned responses in MSTd, showing that selectivity is for global motion pattern. Because most artifacts were eliminated, we were confident that any differential labeling demonstrated for the different patterns could not reflect differential selectivity for a stimulus dimension other than motion pattern.

Recording and Lesioning

Because we expected the motion stimuli to activate several areas within the STS, it was important to be able to distinguish MSTd from nearby cortical areas. One approach we used to roughly delineate this region involved mapping out the relevant area of the STS by single-unit recording. Units located in MSTd were tentatively identified based on their location in the chamber and depth relative to the dura. MSTd and MT were mapped out based on the tuning characteristics of cells in these regions. Particularly helpful in distinguishing MSTd from MT was the former cells' large receptive fields and positional invariance with respect to stimulus placement in the receptive field. After the approximate outlines of MSTd had been established, 100 micro-amp lesions were made around its periphery that could be visualized in histological sections. Only the second of the two monkeys used (macaque 92-2) was explored with these techniques, and recording/lesioning was only performed on the hemisphere that was not subsequently flattened. The flattened hemisphere was left intact, because of concern that mapping this region with micro-electrodes might interfere with the quality of the sectioned tissue.

The details of the awake/behaving recording procedure have been reported elsewhere (Graziano, et al 1994). All recordings were carried out under awake, behaving conditions. A fixation task, described below, was used to maintain a stable retinal image of the stimulus. A scleral search coil and an acrylic skull cap were implanted in the monkey five days before training on the fixation task began. Training and subsequent behavior were reinforced by depriving the monkeys of fluid before each session and then giving drops of apple juice upon correct task completion. Following mastery of the behavior, a craniotomy was performed allowing chronic access to the brain for recording purposes. Around the craniotomy, a plastic chamber was attached to facilitate fixation of a Narashige microdrive, which provided fine control of electrode position. For these experiments, we used varnish-coated tungsten electrodes in conjunction with a guide tube to protect the electrode as it passed through the dura. The chamber was placed vertically for a superior approach to the STS, avoiding striate cortex. Recording sessions lasted 4-6 hours, 5 days a week, typically collecting data from 0-5 cells per day. Each day, the monkey would perform anywhere between 500-2000 trials. The animals were given ample rest periods during each recording session.

The animal was placed 57 cm away from a wide-field, tangent screen, projection monitor, which readily allowed stimuli 40 degrees in diameter to be presented in front of the monkey. Trials were initiated by the appearance of a green (0.1 degree) fixation point directly ahead of the animal. The monkey was required to fixate the target for 500 msec after which the he was exposed to a three second stimulus period during which two complex motion patterns of the type described above were presented. A one second gap, when only the fixation point was visible, was sandwiched in between the two presentations. The center of the stimuli corresponded to the fixation point. Throughout the trial, eye position was monitored. If eye speeds exceeded 15 degrees/sec (as in a saccade), the trial was terminated without a reward. Software monitored eye position every 35 msec and the SD of eye position was always less than 10 min. of an arc. Data collection was controlled by a PC-compatible 486 computer and stimulus presentation was controlled by a PC-compatible 386 computer using a Number Nine board as a display card.

A similar behavioral paradigm was used on the monkey 92-2 during the 2-DG labeling period. However, to maximize the amount of time that the monkey viewed the visual stimulus, the fixation point and motion pattern remained on the screen as long as the monkey maintained fixation. Every 3 seconds of fixation, the monkey received an additional fluid reward. Failure to maintain fixation was followed by a 0.5 second "fixation-off" period before the monkey was allowed to reaquire the target and initiate the next stimulus presentation cycle. Macaque 92-2, the first monkey used, was required to pull a lever in addition to maintaining target fixation. This monkey's behavior was monitored by a PDP-11 computer.

2-DG double labeling

For each of the two labeling periods, a different stimulus was presented to the animal. For the case of macaque 92-1, during the initial 3H-2DG labeling period, a counter-clockwise motion pattern was presented. During the second period (14C-2DG labeling), the monkey viewed an expansion pattern. For the experiment with macaque 92-2, an expansion pattern was used as the first stimulus, followed by contraction during the second period.

The time course of the labeling experiments was a follows. The monkey viewed the first pattern for five minutes before the injection of the first label, to make sure he would consistently perform the behavior before we committed to the procedure. After five minutes, the first label was injected via an IV line and flushed with normal saline. After 45 minutes, the second label was introduced and we changed the stimulus the monkey was viewing. After a 10 minute exposure to this second pattern, the monkey was sacrificed via a pentobarbital overdose and the brain was prepared for storage.

For each of the stimulus periods, the monkey was exposed to the appropriate motion pattern over 85 percent of the 55 minute experimental period. At no point were there significant gaps of time when the monkey would not perform the fixation-task behavior. Post-mortem processing of the specimens involved perfusion with formaldehyde followed by removal of the brain from the cranium. These two steps were completed within five minutes of pentobarbital infusion. The two hemispheres were separated, with one half of the brain frozen intact in dry ice, and the other half was flattened before freezing.

800 times more 3-H isotope was used than 14-C labeled compound, for reasons which will be discussed below under the image processing procedures.

Image Processing

The principle challenge with the double label technique is to recover images from autoradiograms which reflect the label distribution of just one isotope (Friedman et al, 1989; Friedman et al, 1987; Juhler et al, 1987). Obtaining an independent 14-C exposure is relatively simple. Exposing the tissue samples on color film results in the different isotopes activating different layers on the color film. This is because the deeper layers are shielded by the more superficial ones and the low energy beta emissions from the 3-H can only expose the superficial emulsion. By using a yellow filter, which transmits light the same wavelength as the deepest film layer, an isolated image of the 14-C labeling pattern is obtained.

Obtaining a pure 3-H image is technically more difficult. Any film which is sensitive to 3-H emissions is also sensitive to the emissions from 14-C. Many groups have circumvented this problem at the level of the autoradiogram by using a set of radioactive standards along with image subtraction techniques to digitally recover a clean image of the uptake pattern (Friedman et al, 1989; Friedman et al, 1987; Juhler et al, 1987). The approach we decided on was to load the animal with much more 3-H signal than 14-C signal (800:1 ratio) and use special x-ray film sensitive to both isotopes. Because so much more 3-H signal is present in the tissue, this isotope largely overwhelms the 14-C contribution.

The images were scanned into a PC-compatible computer and digitized using standard techniques. The raw data was transferred to a Macintosh recognizable format and the images were edited in Adobe Photoshop. This software was useful for blurring out unwanted tissue artifacts such as rips in the tissue as well as filtering out scratches in the films. These edited images were imported into NIH Image where sequential image slices were averaged together in order to improve the signal to noise ratio. In the unflattened section only two or three images could be pooled in this way, owing to gradual shifting of the columnar pattern in sequentially slices. In the flattened sections, because the cutting was done perpendicular to the columnar pattern, the 2-DG patterns remained in alignment through multiple layers of cortex, allowing large numbers of images to be averaged together.

In regions of the brain where a columnar organization was evident, we were interested in determining the relationship between the labeling patterns visible on the two autoradiograms. This was done by digitally subtracting the gray-level 14-C image from the 3-H image. The resulting images had intermediate density values (appearing gray) in regions of tissue where the two labeling patterns were in register. For this subtraction process to work properly, the two images had to be equated with regards to average pixel luminance and contrast range. Interdigitating white and black columns in the difference image meant that the labeling patterns of the two images were out of phase. An "absolute value" difference image was computed in the same way, except that the pixel values resulting from this process reflected the degree of closeness of the corresponding pixels in the two images. Pixels of the same value were black on the screen. Lighter areas on this image were regions where the pixel values had poor correspondence.

Histology

Selected regions of tissue were subjected to cytochrome oxidase (Seligman et al, 1968; Wong-Riley, 1979) and myelin staining (Gallyas, 1979). Staining patterns obtained using these techniques were digitized and processed using the methods described above for the autoradiograms.

RESULTS

The most striking feature evident on both sets of autoradiograms was the preponderance of labeled columns in several diverse regions of cortex. We found patchy distributions of labeling in auditory cortex, throughout regions of the inferior parietal sulcus, in areas V2 and VP, as well as in the

STS. These columns were oriented orthogonal to the surface of the cortex and extended through most its thickness, although labeling was densest in the middle layers. These columns were generally 0.2 - 0.5 mm in diameter, although the spacing between the columns varied more widely. In general, the columnar organization was better defined on the 14-C images than on the 3-H, which is consistent the 3-H images being somewhat underexposed, and probably does not reflect anything about the underlying distribution of units sensitive to the two patterns.

Although columnar labeling was found in diverse areas of cortex, the uptake patterns for the 14-C and 3-H exposures were closely matched in all regions examined, except for parts of the floor and posterior bank of the STS. The significance of interdigitating versus superimposed columns will be taken up in the discussion. Within the STS, extensive regions of patchy labeling were evident for both sets of films. Figure 2 shows four images of the floor and posterior bank of the STS. Frames A and B show exposure patterns for the 14-C and 3-H isotopes, respectively. Frame C is the difference image obtained from the previous two patterns, while Frame D is the "absolute value" difference image, as described in the METHODS section. These images were obtained from monkey 92-2, where expansion and contraction patterns were compared. From these images it is evident that the region of interdigitation is limited to a fairly well defined area. Beyond this region, patchy labeling continues to be discernible, but the patterns of activity superimpose.

The upper border of the autoradiograms of Figure 2 represents the anterior border of the STS floor. The anterior bank was cut away from this section for purposes of flattening and mounting. Figure 3 shows autoradiograms from anterior bank sections for the two isotopes along with the difference image. The magnification was the same as for the posterior sections, making it apparent that the columns in this bank were less well defined, larger, and relatively well superimposed. Note that the top of Figure 2 and the bottom of Figure 3 were contiguous in the uncut brain, although this is not obvious looking at the two sections because of the way the two halves were dissected. (Because of the angle taken by the blade during cutting, a rim of white matter remained at the parameter of the cut edge.) It is evident that a shift to more diffuse and less well interdigitated columns gradually occurs moving anterior within the floor of the sulcus.

The region of column interdigitation has an interesting topography and does not correspond to the borders of any single anatomical area within the STS. Because of this, distinguishing the border between the different areas was a primary concern. Figure 4 was taken from a paper by Desimone and Ungerleider (1986) where the cortical areas within the STS were mapped for three monkeys using a combination of single-unit recording and histological techniques. Their diagram is in a similar orientation as the tissue in Figure 2, although in their map the anterior bank remains attached. Based on these three maps, it seems likely that the interdigitated area cannot correspond exactly to the borders of MT, although it is likely that area MT partially overlaps this region. The roughly circular patch devoid of label and adjacent to the anterior edge of the interdigitating pattern provides a useful landmark. Based on previously published maps of this region (Figure 4), this area likely represents the far peripheral representation of MT (MTp). Because the motion patterns that the monkey viewed were limited to the central 20 degrees of the visual field, far periphery representations wouldn't be expected to be labeled by either isotope. It is significant to note in Figure 2 that a columnar pattern is absent in this region not just for the difference image, but for the separate isotope images as well. This contrasts with other areas in the STS surrounding the interdigitated section that are labeled by the individual isotopes but which cancel out in the difference image. These latter regions most likely contain units whose receptive fields overlap the stimulus but whose units lack a functional organization along the stimulus dimension sampled by the two pattern types. Given that little 2-DG uptake occurred at all in this oval patch, it is reasonable to propose that this reflects the retinotopy of the region rather than response specificity. Furthermore, it would be highly unlikely to find an island of motion insensitive cells in the STS surrounded by regions giving vigorous responses to these patterns. We believe that this region corresponds to area MTp, an area known to contain units with receptive fields centered in the far periphery and that immediately borders area MT (Ungerleider et al, 1986).

Looking at the three maps of Figure 4, MT should be placed in the area of cortex immediately adjoining the region devoid of label and towards the posterior bank. Based on the Ungerleider sketch of this region, the lateral, medial, and anterior regions bordering the unlabelled tissue would be the various subdivisions of MST. If these assumptions are valid, the difference map showing interdigitated columns in these regions demonstrates a clear functional organization for complex motion pattern in both MT and MST. With respect to the various subdivisions within MST, the region lateral to the unlabelled area could reasonably be assigned MSTl and the region medial, MSTd. Figure 5 shows the same images in Figure 2, with the approximate anatomical borders overlaid. Note that the pattern of interdigitation present in the difference image of Figure 2C undergoes a transition in the region that we have proposed as the MT/MST border. In Figure 2D, this transition is even more dramatic, and reinforces the possibility of multiple anatomical areas included within the region of interdigitation. The part of the interdigitated region labeled MT on the "absolute value" image is distinguished by dark, narrow, winding bands interspersed amongst the lighter patches. The interdigitating pattern within the area we have designated MST lacks these bands. By the convention established in the METHODS section, these darkened areas represent regions of tissue where the labeling pattern was similar for the two isotopes.

Note in Figure 5 that the lateral and posterior borders of MT are incomplete. The most foveal representation of MT is expected to lie along this border, extending beyond the interdigitating region. We would expect the most foveal region of MT to lie outside the area of interdigitation because the portion of the stimulus sampled contains motion not even approximately translational. The functional organization observed in regions of MT somewhat away from the fovea likely reflects a specificity for the approximately linear motion signals towards the outside of the stimuli. Also, because foveal MT receptive fields are small, the number of dots within each receptive field at any time is relatively small and might provide a poor stimulus with which to drive these cells.

According to the borders established in Figure 5, the tissue anterior to these sections (Figure 3) contains areas of MST with more eccentrically placed receptive fields. The weaker labeling and lack of interdigitation observed in the anterior bank sections could be a combination of more peripherally located receptive fields in this region, and therefore weaker stimulation by the foveally placed stimuli, and a gradual loss of functional organization in the anterior regions. The relative contributions of these two effects would be impossible to distinguish without extensive single-unit mapping of this region — a process likely to disrupt the integrity of the tissue and the 2-DG labeling pattern.

Figure 6 shows a horizontal section of the STS that allows visualization of both MT and MST. This tissue came from the right, unflattened hemisphere of macaque 92-2, the same monkey examined in the flattened sections of Figures 2, 3, and 5. Frames A and B show the distributions of 14-C and 3-H label, respectively. Frame C is the difference image obtained from the previous two pictures. Note that the columnar pattern in all three of these images is much more in evidence towards the posterior regions of the sulcus. The posterior bank is at the top of each frame. Significantly, the difference image shows no signs of columnar interdigitation on the anterior bank of the section. The region of differential labeling extends from the posterior bank across most of the extent of the sulcus floor, and ends just posterior to the bend where the anterior bank meets the floor. This is consistent with the labeling patterns obtained in the flattened sections.

We had hoped to make a more definitive statement regarding anatomical borders. In particular, we had hoped to distinguish MT from MST and other neighboring regions based on the greater density of myelin within the former area. Ungerleider and Desimone (1986) have reported that a heavily myelinated region on the floor and posterior bank of the STS corresponds to the foveal projections of V1 and V2 to the STS. Using single unit recording in conjunction with electrolytic lesioning, they identified this area as MT. Injections into peripheral V1 and V2 were transported to a lighter myelin staining region directly adjacent to MT, towards the anterior bank, a region they identified as MTp. In our case, myelin staining of sections cut parallel to the cortical surface of the STS did not adequately demarcate the borders of interest.

The second method we attempted to delineate cortical areas was limited single-unit mapping on the side of the unflattened (right) hemisphere. As described in the METHODS section, lesions were created around the periphery of what we had identified physiologically as MSTd. Even on unstained sections, these lesion points were readily identifiable on the floor and bottom part of the anterior bank of the STS. Unfortunately, because the sections from the two hemispheres were cut completely differently and because of possible asymmetries between hemispheres, little insight into the subdivisions of the flattened sections was achieved. Ideally, the lesions would have been placed in the hemisphere that was flattened, but we were concerned that this would have significantly damaged the tissue.

Based on recent reports that area MT can be distinguished from neighboring regions within the STS based on its cytochrome oxidase levels (Tootell and Taylor, 1995), the tissue was also stained for this enzyme. Unlike myelin, the pattern of cytochrome oxidase was consistent between layers (Figure 7). In addition, the same layers contained both strong deoxyglucose and cytochrome oxidase signals, allowing the two patterns of labeling to be compared. As seen in Figure 7A, the region tentatively identified as MT based on the retinotopic arguments given above, stained darker for this marker than the areas believed to represent MTp and MSTl, lending some support to our boundary assignments. The border between MT and MSTd could not be distinguished using this method, as the tissue in both regions contained high cytochrome oxidase activity. The proposed boundaries are seen most clearly in Figure 7B, where the pixel values of Figure 7A are thresholded.

All the results discussed so far have been for macaque 92-2. The labeling patterns recovered from the other monkey (macaque 92-1) where not nearly as distinct (Figure 8). The images from these samples contained a number of artifacts as a result of tears in the tissue that are visible in the figure. Nevertheless, from the difference image it is clear that the pattern of interdigitation is weak compared with the previous monkey. We suggest that this is a consequence of the stimuli used and reflects the functional organization of the region. The degree of interdigitation is consistent with units driven by expansion and rotation lying closer together than units tuned for expansion and contraction. We would expect that if the two stimuli used were the two directions of rotation, a high degree of interdigitation would have been recovered, similar to that seen for macaque 92-2 (Lagae et al, 1994).

DISCUSSION

We have shown evidence that units tuned to complex motion patterns such as expansion, rotation, and contraction are organized into distinct cortical columns that contain cells of similar tuning along this stimulus dimension. Area MT contains the largest part of the interdigitating region, but we have provided evidence from cytochrome stains and retinotopic maps that some of the columns are present in MST as well. The pattern of interdigitating columns observed in the difference images obtained from the autoradiograms provides evidence that this section of cortex is important in processing information about motion pattern. Since changing only a single stimulus aspect in the two labeling periods resulted in different, anatomically distinct, populations of cells labeled, it is reasonable to conclude this reflects something fundamental about the organization of information processing in the STS.

2-DG autoradiograms must be interpreted carefully. Exposing an animal to a stimulus during a labeling period and then concluding that the pattern of activity on a single autoradiogram reflects specificity for a particular stimulus attribute is unwarranted. For example, in the present study we observed a columnar labeling pattern in regions as diverse as auditory cortex and visual area V2. The columnar organization may reflect local differences in basal metabolic activity, such as that observed in V1 blobs. This may well be the explanation for the strong columnar labeling that we observed in V2 — the similarity between 2-DG patterns and cytochrome oxidase patterns in V2 has been previously noted (Tootell et al., 1983). Alternatively, the labeling pattern may reflect a columnar organization for units tuned to a stimulus dimension other than the one expected. For example, if a hypothetical cortical region with cells arranged into columns based on specificity for stimulus size were exposed to the 40 degree stimuli used in this study, they would demonstrate a periodic labeling pattern.

Using two labels and varying the stimuli for the two exposure periods along only one dimension eliminates these inherent problems. Differential labeling reflecting intrinsic local differences in metabolic activity is eliminated in the difference image, because this contribution to the overall pattern should reasonably be assumed constant over the two stimulus periods. Furthermore, since the stimuli presented during the two time segments differ only in a single quality, labeling patterns reflecting other aspects of the stimuli (size, spatial frequency, color) should also be eliminated after subtracting the two images.

Direction Columns

Several previous investigations by other workers have reported clustering of directionally selective cells in visual cortex. Single unit studies in areas 17 and 18 of the cat have shown cells of similar directional selectivity organized into 300 micron clusters (Swindale et al, 1987; Payne et al, 1981; Tolhurst et al, 1981). At least one group has suggested that this organization is likely to be columnar (Tolhurst et al, 1981) based on detailed mapping with a large number of penetrations. Directional preference tends to change smoothly and gradually moving across primary visual cortex, interrupted by occasional 180 degree reversals in specificity. The presence of frequent 180 degree reversals suggests that (at least in primary visual cortex of the cat) the axis of directionality rather than directional preference per se is organized systematically.

More directly relevant to this study are reports of directional selectivity in the superior temporal sulcus. Based on single-unit mapping of this region, Albright, et al (1984) have reported a clustering of units with similar response selectivity in area MT. Similar to the arrangement in V1 for orientation, Albright found that, moving tangential to the cortical surface, direction preference changes gradually, with periodic 180 degree reversals in specificity. Other workers have made similar observations more informally (Zeki, 1978; van Essen et al, 1981; Maunsell et al, 1983; Baker et al, 1981). Consistent with the occasional reversals in direction preference, Albright proposed that the representation of motion axis in MT is more systematic than direction of motion. He suggested that cells with similar motion axes are organized into continuous slabs of cortex aligned normal to the pial surface, based on the observation that vertical penetrations had the most gradual shifts in axis preference. Albright estimated that 180 degrees of axis representation extends over 400-500 microns of cortex. This size correlates well with the size of orientation and direction columns in area V1.

The current study generally supports Albright's proposed functional organization for MT. Because the stimuli are large compared with the size of an MT neuron's receptive field, the motion is essentially linear in the aperture of space viewed by the unit. Therefore, our stimuli should effectively drive units in MT. The component motion vectors are essentially perpendicular to one another for the case of expansion and rotation and are oriented 180 degrees apart for expansion and contraction. Albright postulates two systems of functional organization within MT. One system alternates with a periodicity of 360 degrees, with opposite directions of motion maximally spatially separated (See Figure 11 of Albright et al, 1984). However, to account for the reversals in stimulus selectivity he observed, some regions of opposite direction preference need to be placed adjacent to one another. This results in a second system organized around axis of directionality, with a 180 degree periodicity. Depending on which of these two systems predominates, a different labeling pattern in the difference image would be predicted. If MT is predominately organized by axis of directionality, expansion/contraction patterns should drive one population of cells and rotation patterns the other population. Therefore, pairing expansion/rotation for the two labeling periods should give the widest possible column separation on the difference image, while expansion/contraction should produce a lesser degree of interdigitation (more overlap) because cells selective for these patterns are less spatially distinct. The images obtained from this study show the opposite pattern of labeling, arguing that the functional anatomy has an organization dominated by the 360 degree, motion direction architecture. Our results are consistent with those obtained in the anaesthetized macaque by Tootell and Born (1991) using translational motion patterns and double-label injections. They found that an interdigitating mosaic of columns was best obtained using patterns oriented in opposite directions.

A point of disagreement between our results and Albright's data is the shape of the isotuning columns. He proposed a "slab-like" arrangement of units with similar selectivity, while our images suggest a cylindrical system of isotuning columns. A possible way of reconciling the 2-DG and single-unit recording data takes into account the band vs. interband distinction reported in MT (Born et al, 1992). During a single-label 2-DG experiment, if a monkey is exposed, during a single labeling period, to wide-field patterns sampling a full range of motion directions, it might be predicted that area MT would be homogeneously labeled. Surprisingly, the distribution of label is patchy, suggesting a system of columnar organization distinct from that organized around directionality (Born et al, 1992). The areas of high 2-DG uptake were called "band" regions and the surrounding regions designated "interband." It was determined that the units in interband regions were directionally selective but contained inhibitory surrounds and therefore responded poorly to the wide-field motion patterns presented during the labeling period. Because large motion stimuli were used in the current study, the pattern of labeling observed likely reflects the presence of direction columns only within band regions. We should point out that the autoradiogram patterns cannot be explained purely on the band/interband distinction alone because without the directional organization the difference images would show a homogeneous gray rather than the observed interdigitating mosaic. However, the presence of band/interband regions superimposed onto Albright's proposed iso-tuning slabs could result in the cylindrical labeling pattern recorded in our images.

Functional Organization of MST

Despite taking several approaches to delineate the MST/MT border, we were unable to obtain a completely satisfying map of this region. Figure 5 shows our best attempt based on previous reports about the retinotopic organization of this area. Although the data from the myelin staining was noninformative, the pattern recovered from the cytochrome oxidase sections had similarities with the 2-DG images. Most strikingly, the region which we designated MTp based on retinotopic arguments turned out to be a cytochrome oxidase poor area. Why an anatomical region representing the peripheral part of the visual field should be relatively deficient in this enzyme is not clear, but the fact that MTp can be potentially distinguished from surrounding regions based on this marker provides an additional argument for considering this region anatomically distinct from MT. As a proposal for future studies, making an HRP injection into peripheral V1/V2 should result in axonal transport of this enzyme into the region we have tentatively called MTp.

We have placed MSTl on the strip of cortex just lateral to MTp, consistent with its location in other studies. The receptive field size and tuning properties of units in MSTl for motion stimuli are similar to those found in MT; the two areas differ physiologically largely with respect to their responses during smooth pursuit (Wurtz et al, 1988; Newsome et al, 1988; Komatsu et al, 1988). Like neurons in MT, MSTl cells view only a small portion of the large patterns used in this investigation and therefore many are being driven largely by a translational motion signal. Both regions also contain a subpopulation of neurons with inhibitory surrounds, although whether MSTl contains an analogous band/interband organization is not known. The pattern of MSTl interdigitation evident in the difference images possibly reflects both directional and band/interband functional organization. The pattern of cytochrome oxidase staining is consistent with this strip being anatomically distinct from MT — enzymatic activity drops off quickly at the border between these two regions, although levels remain higher than in neighboring MTp. It is interesting to note that the frequency content of the labeling pattern appears to change at what we have assigned the MT/MSTl border, offering further support for the distinctness of these two regions.

The case for placing MSTd in the region just medial to MTp is less convincing. Like the adjoining tissue in MT, the cytochrome oxidase activity is high along this strip. The pattern of interdigitation is somewhat different than that observed in MT, but this is not particularly compelling. It is known that units in this area contain cells with large receptive fields that would allow them to analyze wide-field complex motion patterns. Although it has been well established that many units in MSTd are specifically tuned for complex motion patterns, many neurons in this region are selective for wide-field linear motion stimuli, either in isolation or in combination with a selectivity for a complex motion pattern (Duffy et al, 1991; Graziano et al, 1994). From physiology experiments, it is known that units in MSTd tuned exclusively for linear motion can be driven by complex motion patterns, although they are not specifically tuned for these patterns. Because of this, the pattern of labeling observed in MSTd would be expected to reflect not just specificity for complex motion patterns, but also inadvertent labeling of units responding to the component translational motion signals from which these more complex patterns are built. Single unit mapping studies do not exhibit this limitation, because tuning specificity can be established by moving the stimulus pattern around within the receptive field of the unit (Lagae et al, 1994).

Significance of cortical columns

The images obtained in this study lend further support to the belief that mammalian cortex is organized into anatomical columns that are composed of functionally discrete subunits (for a review, see Tootell et al, 1993). We saw evidence of columnar organization throughout diverse regions of the brain, including auditory cortex, somatosensory cortex, and visual cortex, although we found a systematic representation of motion pattern only on the posterior bank and floor of the superior temporal sulcus. We observed no such pattern in area V1, most likely because the dot densities in our stimuli were too low to effectively stimulate the small receptive fields in this region.

Finding a systematic representation of motion direction and complex motion pattern in the STS argues that these visual qualities are extracted by the nervous system at a fundamental level of image processing. The presence of units specifically tuned for these stimulus qualities has already been established, but seeing a well-ordered functional organization provides additional evidence that the nervous system analyzes the visual world along these stimulus dimensions.

What is gained by building the cortex from functional subunits containing neurons of similar specificity? One possibility involves efficiency of establishing connections between different neurons. If it is valid to assume that neurons with similar response profiles and receptive field locations make similar types of connections, by spatially clustering these regions much less information is required to specify the correct connectivity pattern. Clustering changes the scale at which connections need to be specified, making the coding coarser. Establishing and specifying connections in terms of neuronal clusters is obviously more efficient than having to organize these connections at the scale of individual neurons. Furthermore, the accuracy of the projections is less important. A neuron arising from a cell body in one area of cortex need only to find its way with a precision comparable to the size of a cortical column. In addition, since selectivity appears to change gradually across neighboring columns, even greater tolerance in accuracy of connectivity is allowable. A neuron trying to find its way to an orientation column in V1 containing units tuned to vertical contours that misses by a small amount will connect with neurons of similar selectivity.

Mitchison (1992) has argued that dividing the brain up into different anatomic regions and further breaking up these regions into discrete functional subunits minimizes wiring volume. Brain size is likely an important variable to minimize for both reasons of metabolic efficiency and portability.

Another reason for maintaining a systematic organization according to unit specificity is that anatomical spatial relationships offer a dimension for encoding information. As an analogy, the meaning of a word in a sentence is partially determined by its position. Imposing order through a columnar organization is a way of bundling similar types of information together — like a well conceived filing system. From the point of view of establishing representations of the world in a manner that is readily decodable by other brain regions, such a well-ordered structuring of selectivity is clearly advantageous. To see why this is the case, consider the task of an observer attempting to reconstruct an organism's perceptual experience based on viewing the pattern of cortical activity. This task is obviously facilitated by having units of similar specificity clustered together.

Besides enabling efficient specification of connectivity, clustering units of similar specificity together also facilitates the formation of local circuits via the dendritic processes of projection neurons and the processing provided by local interneurons (Tootell et al, 1993). The fact that the size of a typical cortical column corresponds well with the horizontal branching expanse of a projection neuron is unlikely to be coincidental. This formulation of a cortical column as a partially self-contained information processing module has considerable intuitive appeal. Clustering together neurons that are functionally related allows for the shortest possible connection paths and the greatest processing speeds.

Finally, Shadlen and Newsome (1994) have argued that the intrinsic noisiness of a single neuron's rate code requires redundancy to smooth out this variability. Clustering neurons of similar specificity into cortical columns is an efficiency way of introducing this redundancy without further increasing the complexity of the wiring.

Final Comments

To our knowledge this is the first time a double-label 2-DG experiment has been performed on an awake, behaving animal. The advantages of using conscious organisms for 2-DG studies are numerous. First, there is always the concern that the response profiles of cells may change under anesthesia, making generalizations to the awake animal tenuous. Specifically, cortical activity has been shown to be surpressed under the influence of anesthesia and the metabolic rate throughout the gray matter becomes more uniform (Sokoloff, 1977; Sokoloff et al, 1977). The resulting decrease in signal/noise ratio is obviously not desired, especially in studies where this is already a problem.

However, the disadvantages of using awake specimens are not insignificant. Because the animal must perform at least a fixation behavior, the possibility for non-compliance on the day of the labeling is a real concern. Once an animal is injected with radioactive tracer, failure to reliably perform the behavior necessitates sacrificing the organism without achieving meaningful labeling. In addition, the presence of a conscious, incompletely immobilized, monkey with radioactive label requires special handling procedures. Lastly, extra time is required to train the animals on the behavior to the point where performance is reliable for a one hour period.

The advantages of the double-label method over experiments with a single label have been enumerated above and in other studies (Friedman et al, 1989; Redies et al, 1987). Briefly, using two isotopes allows each animal to serve as its own control. The main challenge with these experiments is in recovering pure images that are specific for one isotope or the other, by either exploiting differences in isotope half-life (18-F vs. 3-H) (Redies et al, 1987; Schoppmann et al, 1981; Redies et al, 1989; Friedman et al 1989; John et al, 1986) or in the energy of their emissions (3-H vs. 14-C) (Redies et al, 1990; Altenau et al, 1978). As discussed in METHODS, the choice of 3-H and 14-C results in unwanted cross-contamination of the 3-H film by 14-C activity on the unshielded x-ray film. Our approach was to use a 800:1 ratio of 3-H to 14-C. The 3-H signal effectively overwhelmed the 14-C contribution because its relative activity was so much greater. We believe that this method is cleaner and less time intensive than other solutions to this contamination problem, which have included image processing techniques in conjunction with the use of radioactive standards.

In conclusion, using the 2-DG double labeling technique we have demonstrated the presence of a columnar organization in areas of cortex sensitive to complex motion patterns such as expansion, rotation and contraction. We were able to visualize this organization in terms of a difference image showing alternating peaks in label activity. This interdigitation was more perfect for expansion/contraction than expansion/rotation, suggesting that units driven by expansion and contraction patterns are maximally separated across the surface of the cortex. Unfortunately, the exact borders of the regions on the floor of the STS could not be assigned with complete confidence based on the anatomical methods we employed. We have argued, based on retinotopy, cytochrome labeling, and pattern of periodic labeling evident on difference images, that the area of interdigitation partially overlaps both regions MT and MST.

REFERENCES

Albright, T. D. (1984). Direction and Orientation Selectivity of Neurons in Visual Area MT of the Macaque. Journal of Neurophysiology, 52, 1106-1130.

Albright, T. D. (1993). Cortical processing of visual motion. [Review]. Reviews of Oculomotor Research, 5177-201.

Albright, T., Desimone, R. & Gross, C. (1984). Columnar Organization of Directionally Selective Cells in Visual Area MT of the Macaque. Journal of Neurophysiology, 51, 16-31.

Albus, K. (1979). 14C-deoxyglucose mapping of orientation subunits in the cats visual cortical areas. Experimental Brain Research, 37, 609-13.

Albus, K. & Sieber, B. (1984). On the spatial arrangement of iso-orientation bands in the cat's visual cortical areas 17 and 18: a 14C-deoxyglucose study. Experimental Brain Research, 56, 384-8.

Altenau, L. L. & Agranoff, B. W. (1978). A sequential double-label 2-deoxyglucose method for measuring regional cerebral metabolism. Brain Research, 153, 375-81.

Baker, J. F., Peterson, S. E., Newsome, W. T. & Allman, J. M. (1981) Visual response properties of neurons in four extrastriate visual areas of the owl monkey (Aotus trivirgatus): a quantitative comparison of medial, dorsomedial, dorsolateral, and middle temporal areas. Journal of Neurophysiology, 45, 397-416.

Bartfeld, E. & Grinvald, A. (1992). Relationships between orientation-preference pinwheels, cytochrome oxidase blobs, and ocular-dominance columns in primate striate cortex. Proceedings of the National Academy of Sciences of the United States of America, 89, 11905-9.

Berman, N. E., Wilkes, M. E. & Payne, B. R. (1987). Organization of orientation and direction selectivity in areas 17 and 18 of cat cerebral cortex. Journal of Neurophysiology, 58, 676-99.

Blasdel, G. G. (1992). Orientation selectivity, preference, and continuity in monkey striate cortex. Journal of Neuroscience, 12, 3139-61.

Born, R. & Tootell, R. (1992). Segregation of global and local motion processing in primate middle temporal visual area. Nature, 357, 497-499.

Desimone, R. & Ungerlieder, L. G. (1986). Multiple Visual Areas in the Caudal Superior Temporal Sulcus of the Macaque. The Journal of Comparative Neurology, 248, 164-189.

DeYoe, E. A. & Van Essen, D. C. (1988). Concurrent processing streams in monkey visual cortex. Trends in Neuroscience, 11, 219-226.

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 Neurophysiology, 65, 1346-1359.

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 Neurophysiology, 65, 1329-1345.

Friedman, H. R., Bruce, C. J. & Goldman-Rakic, P. S. (1989). Resolution of metabolic columns by a double-label 2-DG technique: interdigitation and coincidence in visual cortical areas of the same monkey. Journal of Neuroscience, 9, 4111-21.

Friedman, H. R., Bruce, C. J. & Goldman-Rakic, P. S. (1987). A sequential double-label 14C- and 3H-2-DG technique: validation by double-dissociation of functional states. Experimental Brain Research, 66, 543-54.

Gallyas, F. (1979). Silver staining of myelin by means of physical development. Neurological Research, 1, 203-209.

Graziano, M. S., Andersen, R. A. & Snowden, R. J. (1994). Tuning of MST neurons to spiral motions. Journal of Neuroscience, 14, 54-67.

Hubel, D. H. & Wiesel, T. N. (1962). Receptive fields, binocular interaction, and functional architecture in the cat's visual cortex. Journal of Physiology, 165: 106-154.

Humphrey, A. L. & Norton, T. T. (1980). Topographic organization of the orientation column system in the striate cortex of the tree shrew (Tupaia glis). I. Microelectrode recording. Journal of Comparative Neurology, 192, 531-47.

John, E. R., Tang, Y., Brill, A. B., Young, R. & Ono, K. (1986). Double-labeled metabolic maps of memory. Science, 233, 1167-75.

Juhler, M. & Diemer, N. H. (1987). A method for 14C and 3H double-label autoradiography. Journal of Cerebral Blood Flow & Metabolism, 7, 572-7.

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.

Lagae, L., Maes, H., Raiguel, S., Xiao, D. K. & Orban, G. A. (1994). Responses of macaque STS neurons to optic flow components: a comparison of areas MT and MST. Journal of Neurophysiology, 71, 1597-626.

LeVay, S., Hubel, D. H., Wiesel, T. N. (1975). The pattern of ocular dominance columns in macaque visual cortex revealed by a reduced silver stain. Journal of Comparative Neurology. 159: 559-576.

Maunsell, J. & van Essen, D. C. (1983). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. Journal of Neurophysiology, 49, 1127-1147.

Mitchison, G. (1992). Axonal trees and cortical architecture. Trends in Neuroscience, 15: 122-126.

Mountcastle, V. B. (1957). Modality and topographic properties of single neurons of cat's somatic sensory cortex. Journal of Neurophysiology, 20, 408-434.

Nawrot, M. & Blake, R. (1989). Neural Integration of Information Specifying Structure from Stereopsis and Motion. Science, 244, 716-718.

Newsome, W. T., Wurtz, R. H. & Komatsu, H. (1988). Relation of Cortical areas MT and MST to pursuit eye movements. II. Differentiation of retinal from extraretinal inputs. Journal of Neurophysiology, 60, 604-620.

Olds, J. L., Frey, K. A., Ehrenkaufer, R. L. & Agranoff, B. W. (1985). A sequential double-label autoradiographic method that quantifies altered rates of regional glucose metabolism. Brain Research, 361, 217-24.

O'Neill, W. E., Frisina, R. D. & Gooler, D. M. (1989). Functional organization of mustached bat inferior colliculus: I. Representation of FM frequency bands important for target ranging revealed by 14C-2-deoxyglucose autoradiography and single unit mapping. Journal of Comparative Neurology, 284, 60-84.

Payne, B. R., Berman, N. & Murphy, E. H. (1981). Organization of direction preferences in cat visual cortex. Brain Research, 211, 445-50.

Raiguel, L. & Orban, G. (1993). Speed and Directional Selectivity of Macaque Middle Temporal Neurons. Journal of Neurophysiology, 69, 19-39.

Redies, C. D. ,. (1989). Tracer Kinetics Constants and Simplified versions of the double-label deoxyglucose method. Neuroscience, 30, 558-561.

Redies, C., Diksic, M., Evans, A. C., Gjedde, A. & Yamamoto, Y. L. (1987). Double-label autoradiographic deoxyglucose method for sequential measurement of regional cerebral glucose utilization [see comments]. Neuroscience, 22, 601-19.

Redies, C., Diksic, M. & Riml, H. (1990). Functional organization in the ferret visual cortex: a double-label 2-deoxyglucose study. Journal of Neuroscience, 10, 2791-803.

Redies, C., Diksic, M. & Rimi, H. (1990). Functional Organization in the Ferret Visual Cortex: A Double-Label 2-Deoxyglucose Study. Journal of Neuroscience, 10, 2791-2803.

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, 145-57.

Schoppmann, A. & Stryker, M. P. (1981). Physiological evidence that the 2-deoxyglucose method reveals orientation columns in cat visual cortex. Nature, 293, 574-576.

Seligman, A. M., Karnovsky, M. J., Wasserkrug, H. L., Hanker, J. S. (1968). Nondroplet ultrastructure of cytochrome oxidase activity with a polymerizing osmiophilic reagent, diaminobenzidine (DAB). The Journal of Cell Biology, 38, 1-13.

Serviere, J., Webster, W. R. & Calford, M. B. (1984). Isofrequency labelling revealed by a combined [14C]-2-deoxyglucose, electrophysiological, and horseradish peroxidase study of the inferior colliculus of the cat. Journal of Comparative Neurology, 228, 463-77.

Shadlen, M. N. & Newsome, W. T. (1994). Noise, neural codes and cortical organization. Current Opinions in Neurobiology, 4:569-579.

Singer, W. (1981). Topographic organization of orientation columns in the cat visual cortex. A deoxyglucose study. Experimental Brain Research, 44, 431-6.

Skeen, L. C. (1977). Odor-induced patterns of deoxyglucose consumption in the olfactory bulb of the tree shrew, Tupaia glis. Brain Research, 124, 147-53.

Sokoloff, L. (1977). Relation between physiological function and energy metabolism in the central nervous system. [Review]. Journal of Neurochemistry, 29,

Sokoloff, L., Reivich, M., Kennedy, C., Des Rosiers, M. H., Patlak, C. S., Pettigrew, K. D., Sakurada, O. & Shinohara, M. (1977). The [14C]deoxyglucose method for the measurement of local cerebral glucose utilization: theory, procedure, and normal values in the conscious and anesthetized albino rat. Journal of Neurochemistry, 28, 897-916.

Swindale, N. V., Matsubara, J. A. & Cynader, M. S. (1987). Surface organization of orientation and direction selectivity in cat area 18. Journal of Neuroscience, 7, 1414-27.

Tanaka, K., Sugita, Y., Moriya, M. & Saito, H. (1993). Analysis of Object Motion in the Ventral Part of the Medial Superior Temporal Area of the Macaque Visual Cortex. Journal of Neurophysiology, 69, 128-142.

Tanaka, K., Fukada, Y. & Saito, H. (1989). Underlying Mechanisms of the Response Specificity of Expansion/Contraction and Rotation Cells in the Dorsal Part of the Medial Superior Temporal Area of the Macaque Monkey. Journal of Neurophysiology, 42, 642-656.

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.

Tootell, R. B. H. & Born, R. T. (1991). Architecture of primate area MT. Society of Neuroscience Abstracts, 17:524.

Tootell, R. B. H., Born R. T., Ash-Bernal, R. (1993) Columnar Organization in Visual Cortex in Non-Human Primates and Man. From Functional Organization of the Human Visual System, Pergamon Press Ltd.

Tootell, R. B. H., Taylor, J. B. (1995). Anatomical evidence for MT and additional cortical visual areas in humans. Cerebral Cortex, 1:39-55.

Theurich, M., Muller, C. M. & Scheich, H. (1984). 2-Deoxyglucose accumulation parallels extracellularly recorded spike activity in the avian auditory neostriatum. Brain Research, 322, 157-61.

Tolhurst, D. J., Dean, A. F. & Thompson, I. D. (1981). Preferred direction of movement as an element in the organization of cat visual cortex. Experimental Brain Research, 44, 340-2.

Ungerleider, L. & Desimone, R. (1986). Cortical Connections of Visual Area MT in the Macaque. Journal of Comparative Neurology, 248, 190-222.

Ungerleider, L. G. & Desimone, R. (1986). Projections to the superior temporal sulcus from the central and peripheral field representations of V1 and V2. Journal of Comparative Neurology, 248, 147-163.

van Essen, D. C., Maunsell, J. H. R., Bixby, J. L. (1981) The middle temporal visual area in the macaque: myeloarchitecture, connections, functional properties and topographical connections. Journal of Comparative Neurology, 199, 293-326.

Wong-Riley, M. Changes in the visual system of monocularly sutured or enucleated cats demonstrated with cytochrome oxidase histochemistry. (1979) Brain Research, 171, 11-28.

Wurtz, R. H. & Komatsu, H. (1988). Relation of cortical area MT and MST to pursuit eye movements. I. Localization and visual properties of neurons. Journal of Neurophysiology, 60, 580-603.

Zeki, S. M. (1978). Functional specialization in the visual cortex of rhesus monkey. Nature, 274, 245-272.

LEGENDS

Figure 1: Stimuli used in this study. Random dot patterns of rotation (A), expansions (B), and contraction (C). Monkey 92-1 viewed pattern A during the first labeling period and pattern B during the second. Monkey 92-2 viewed pattern B for the first labeling period followed by pattern C. Details of how these stimuli were constructed are provided in the text.

Figure 2: Autoradiograms and difference images of the posterior bank and floor of the STS. Frame A shows the distribution of the 14-C label, Frame B the 3-H label. Frame C is the difference image and Frame D is the "absolute value" difference image.

Figure 3: Autoradiograms and difference images of the STS anterior bank. Frame A shows the distribution of the 14-C label, Frame B the 3-H label. Frame C is the difference image.

Figure 4: Maps of the STS showing multiple cortical areas. This diagram is taken from Figure 1 of Desimone & Ungerleider (1986). Their maps were constructed using a combination of single unit recording and myelin staining.

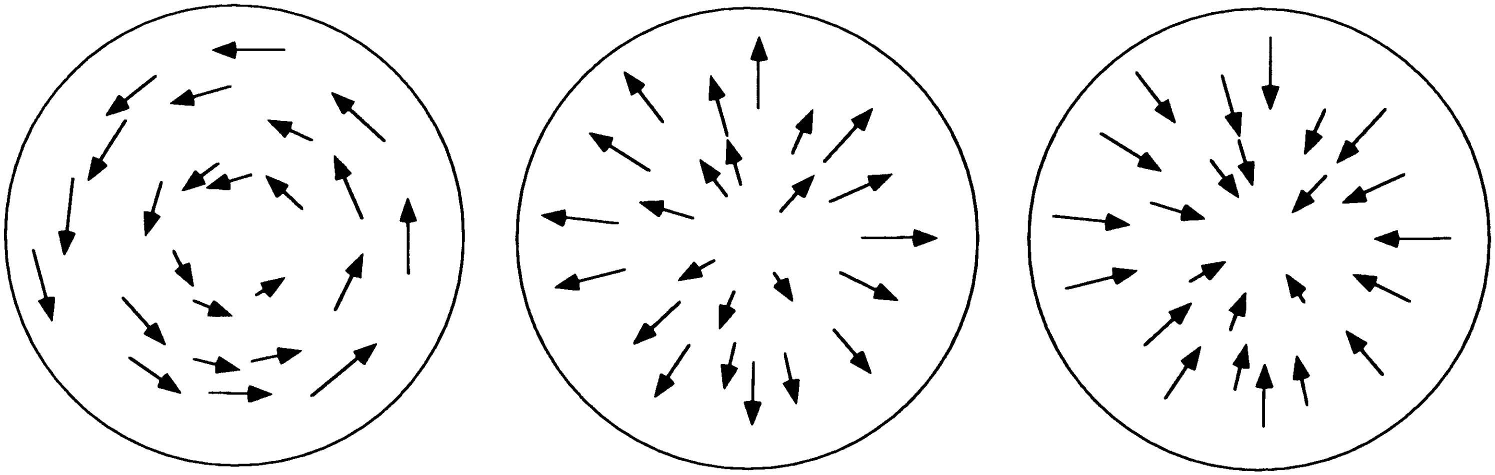
Figure 5: Same images as in Figure 2, but with proposed anatomical borders overlaid. See text for details on how these maps were constructed.

Figure 6: Cross sections of the STS, including areas MT and MST. Autoradiograms come from the unflattened (right) hemisphere of macaque 92-2. "A" shows the exposure for the 3-H label, "B" for the 14-C label, and Frame "C" is the difference image. The posterior bank is towards the top of each frame. Note on the difference image that the region of interdigitation ends just before the anterior bank.

Figure 7: Cytochrome stain of the STS floor and posterior bank. Section shown matches the autoradiograms of Figures 2 & 5. Frame "A" shows the gray-level representation of the staining pattern. Frame "B" is equivalent to the previous image, except that the pixel intensities have been thresholded to help establish borders more clearly. Frame "C" shows cytochrome staining pattern with the proposed outlines of the relevant cortical regions from Figure 5.

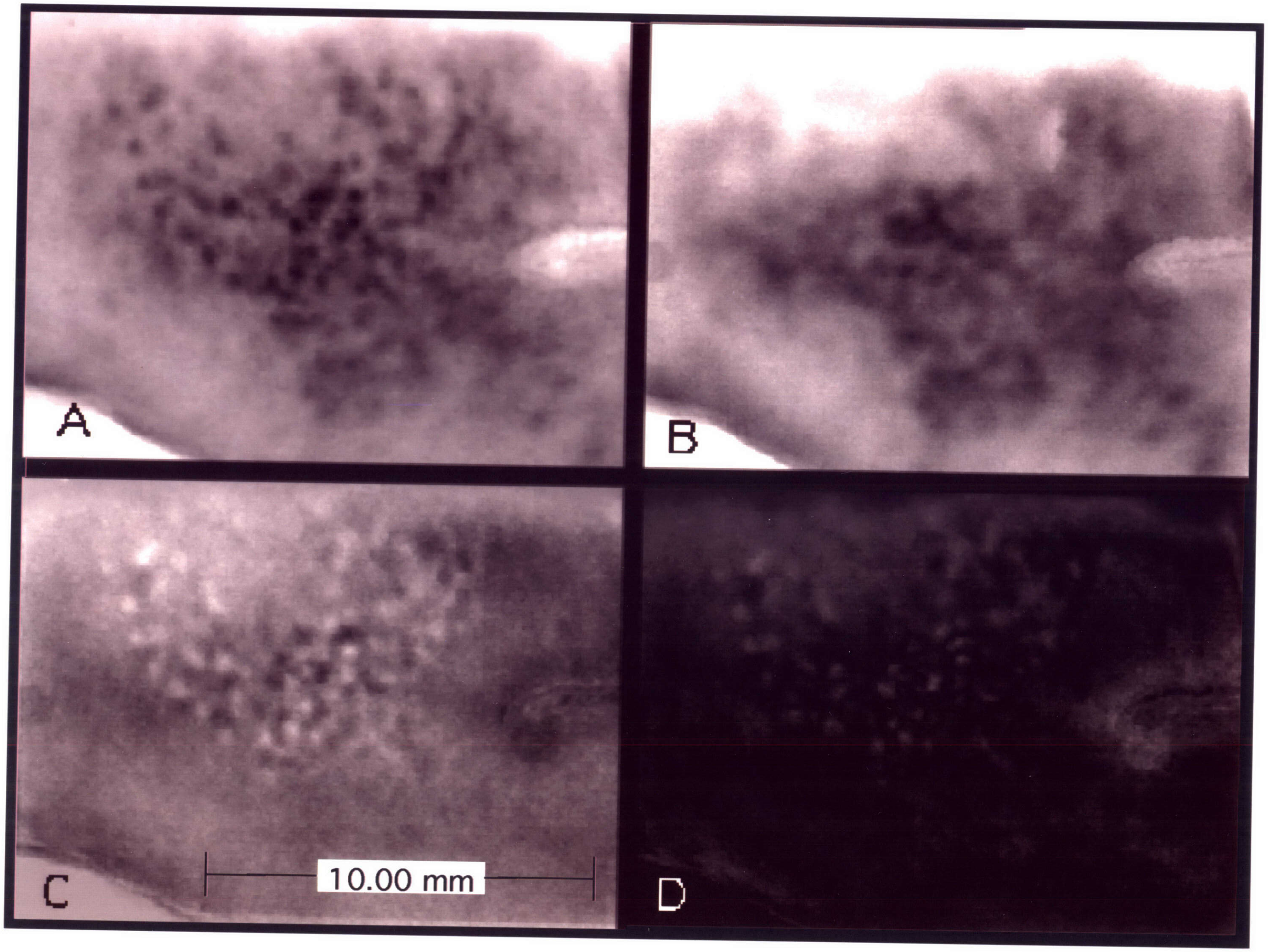
Figure 8: Autoradiograms of the STS floor and posterior bank. Sections are from macaque 92-1, where expansion and rotation were compared. Frame A shows the 14-C labeling pattern, Frame B the labeling pattern for 3-H. Frame C is the difference image. Note that the pattern of interdigitation is weak for reasons discussed in the text. A large tear in the tissue partially distorts all three images.

Figure 1



Rotation Expansion Contraction

Figure 2



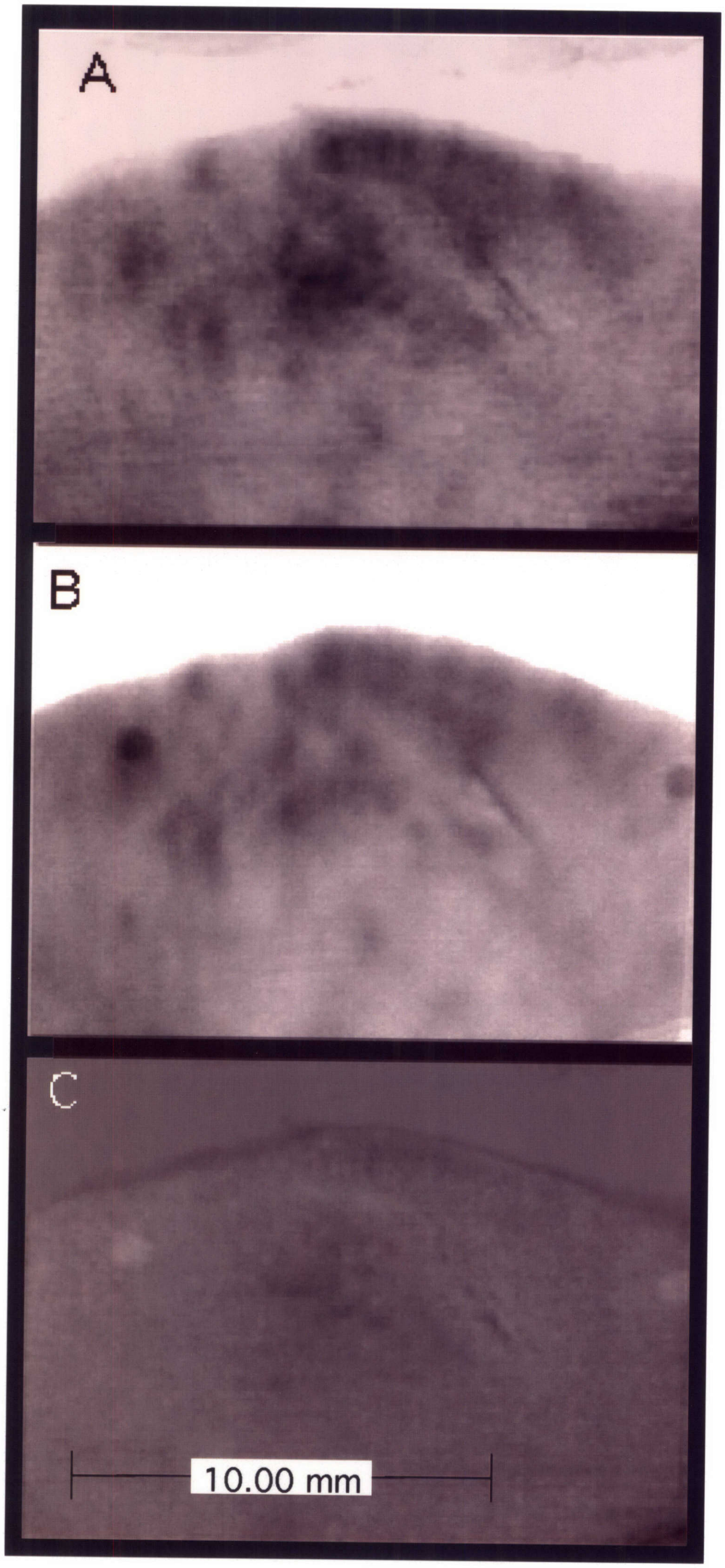
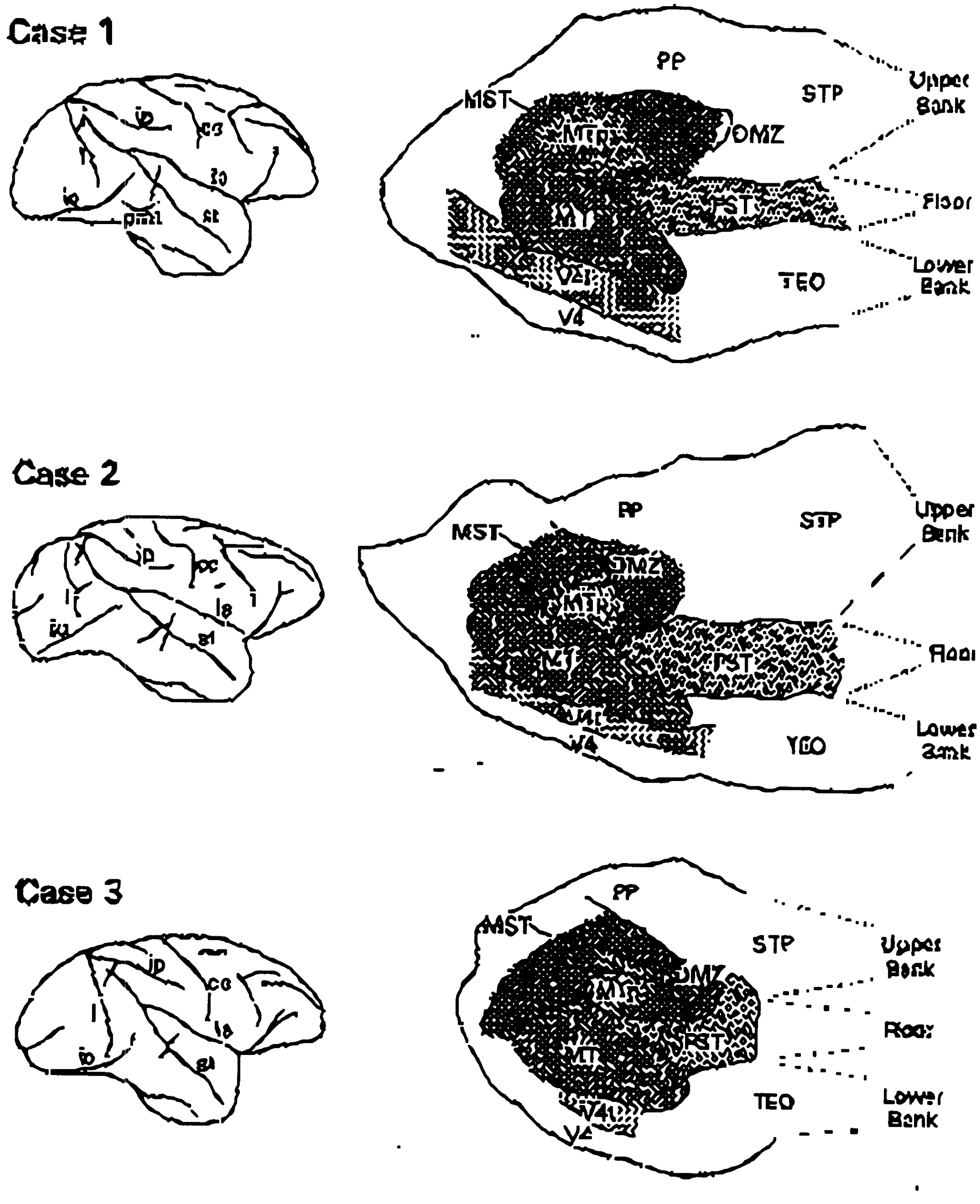


Figure 3

Figure 4


Case 1

Case 2

Case 3

Figure 5

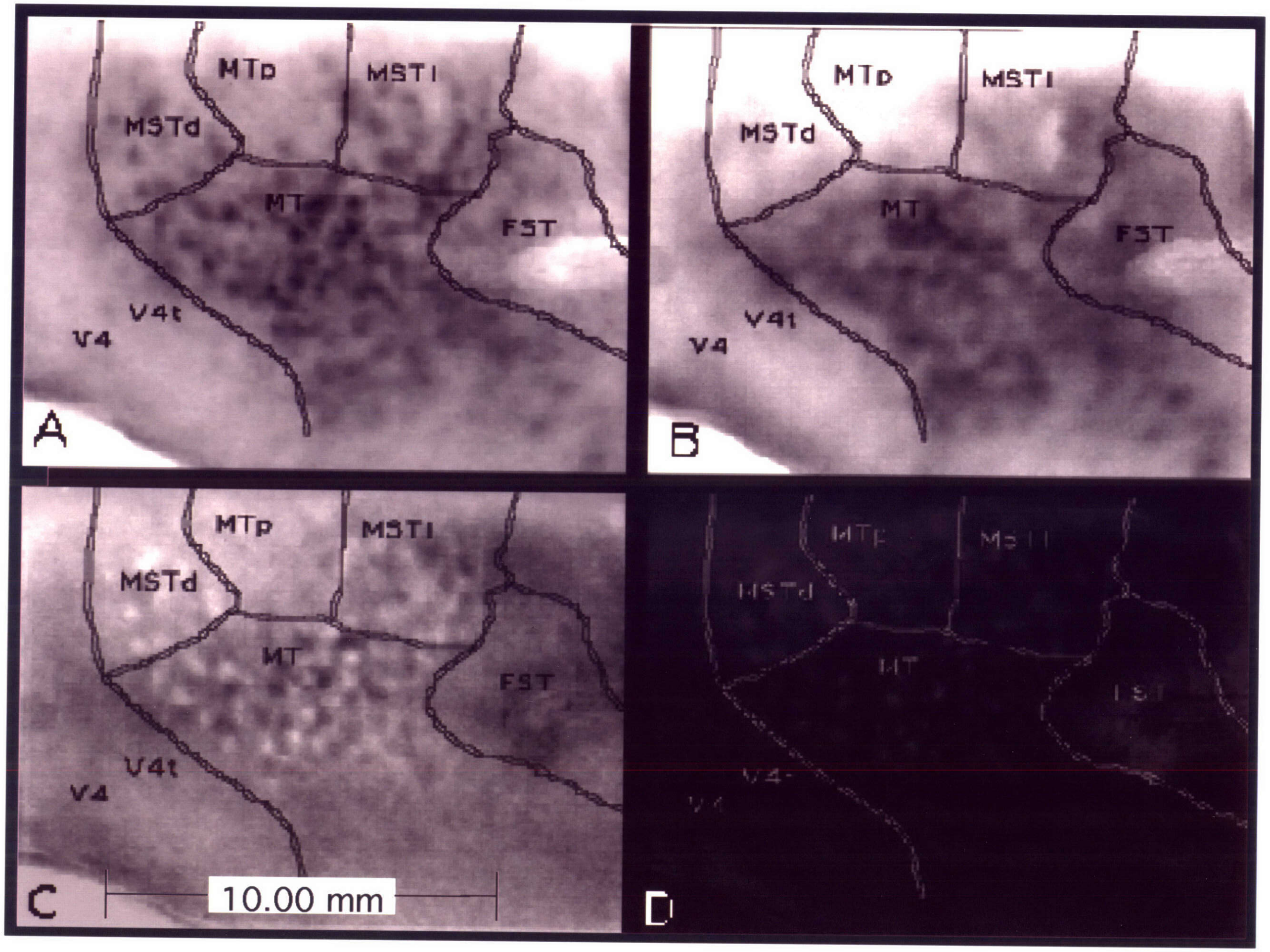
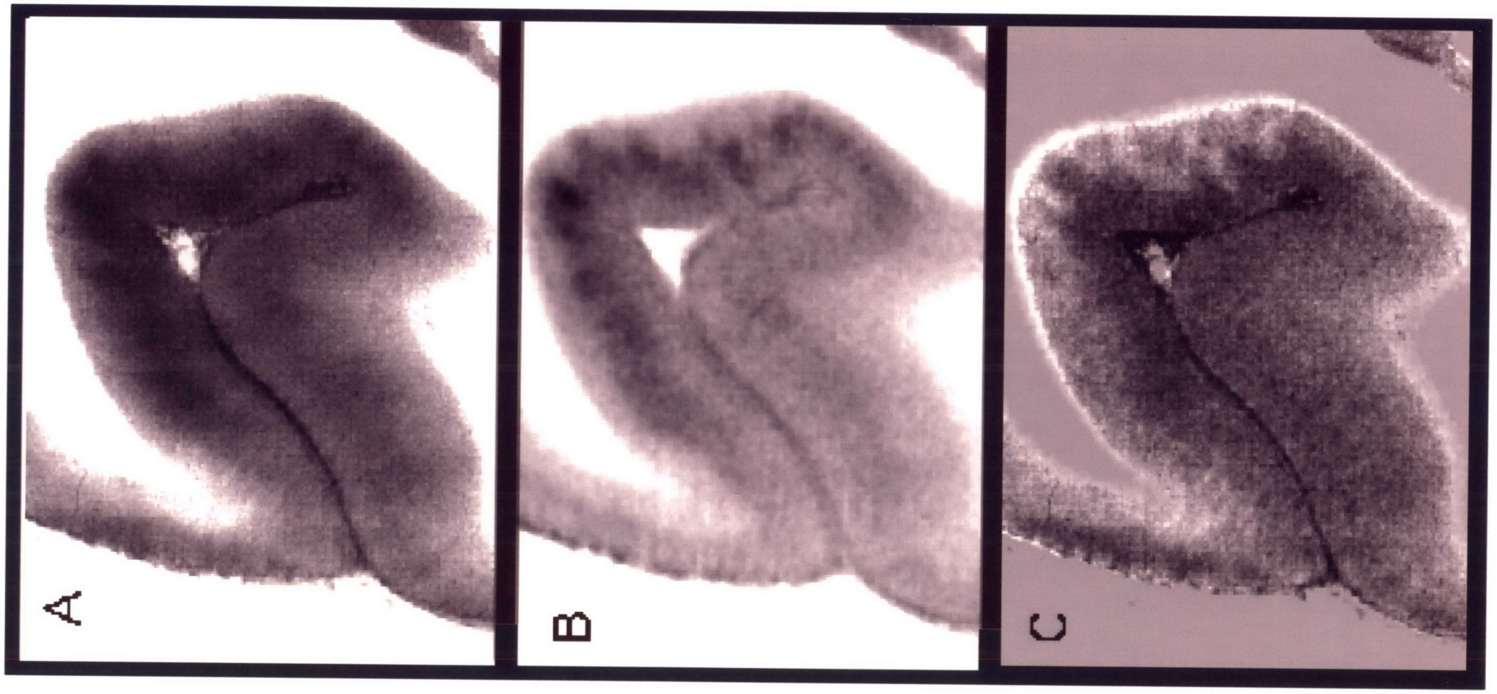

A
MTp
MSTl
MSTd
MT
FST
V4t
V4
B
MTp
MSTl
MSTd
MT
FST
V4t
V4
C
MTp
MSTl
MSTd
MT
FST
V4t
V4
10.00 mm
D
MTp
MSTl
MSTd
MT
FST
V4t
V4

Figure 6



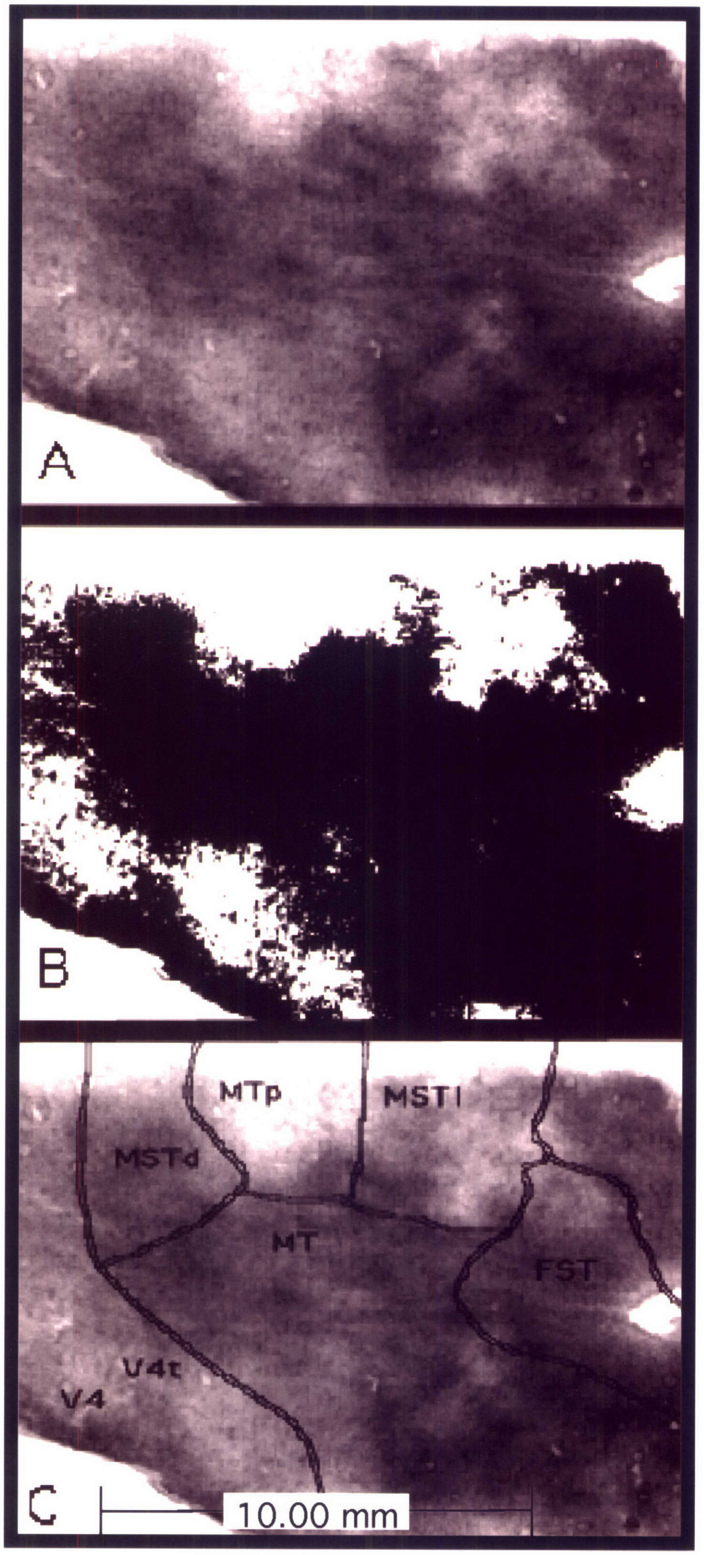


Figure 7

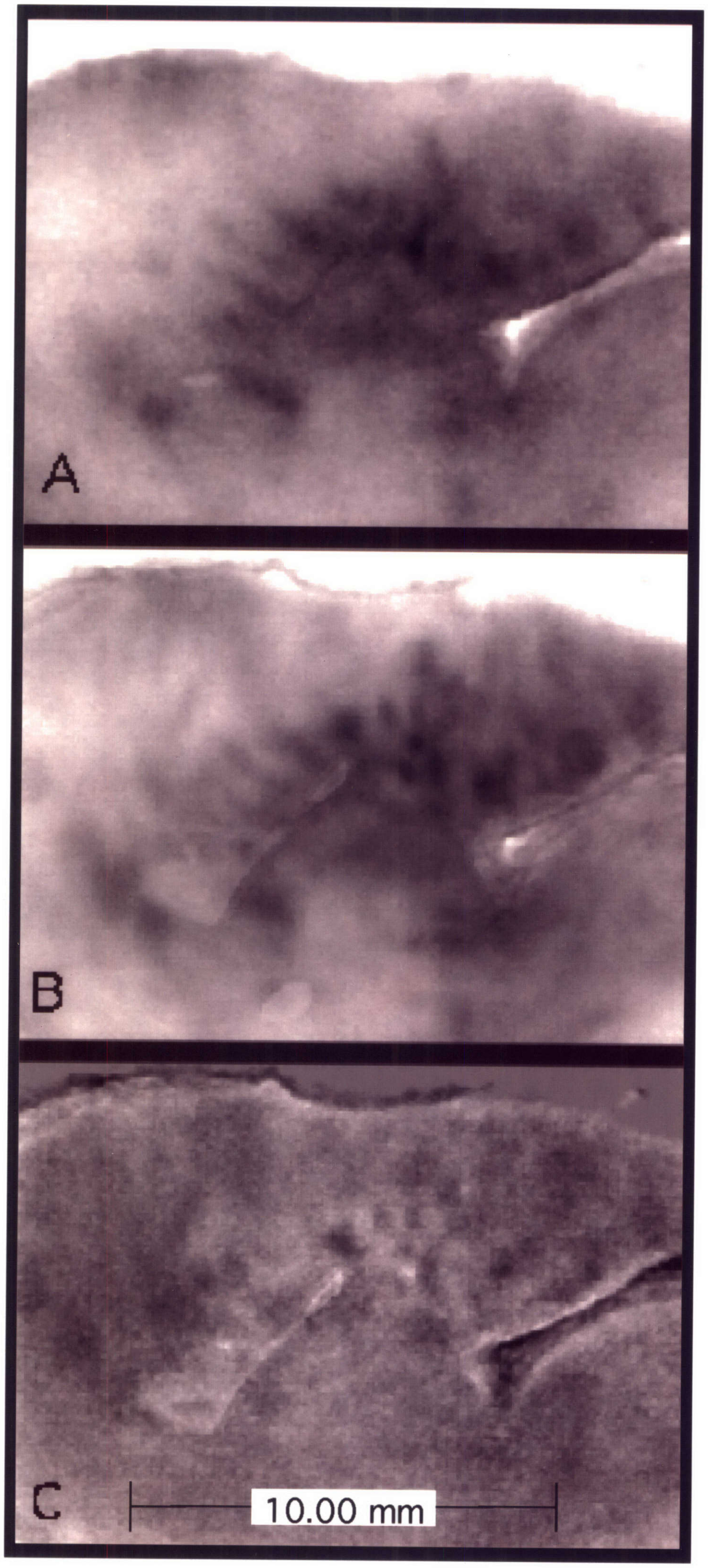


Figure 8

Chapter 4

A Novel Speed Illusion Involving Complex Motion Patterns Related to Cortical Area MSTd

ABSTRACT

Single unit recording studies in area MSTd of the macaque have reported the existence of cells tuned to complex motion patterns such as expansion, rotation and contraction. By a factor of about 2:1, more cells with selectivity for expanding patterns have been found than those tuned for either direction of rotation. We wanted to see if this anisotropy of response selectivity had perceptual consequences, assuming that the same asymmetry in tuning distribution was present in human subjects. Using random dot stimuli well-controlled in terms of the speeds of the features defining these motion patterns, we found that when local motion vectors are organized into a global motion pattern consistent with expansion, their perceived speeds appear greater than when organized into a rotation pattern. In displays where the number of motion directions used to define the patterns was reduced, the magnitude of the illusion decreased. Similarly, the strength of the effect diminished as dot density was reduced. In patterns where only wedge-shaped segments of the stimuli were left exposed, the difference in perceived speed between expansion and rotation increased with the angular extent of the stimulus. Stimulus placement relative to the fixation point had little effect on the persistence of this phenomenon -- expansion patterns appeared to contain elements of greater speed, independent of stimulus eccentricity. These results correlated well with the response profiles of units in MSTd and argue against a local explanation for this perceptual illusion, suggesting that the global motion pattern of the stimulus, per se, is responsible. We hypothesize that the illusion is due to the disproportionate number of expansion cells in MSTd, and the results are discussed in the context of speed perception models. To our knowledge, this is the most direct demonstration to date of a perceptual phenomenom linked to MSTd physiology.

INTRODUCTION

Cells in the dorsal part of the medial superior temporal region (MSTd) of the macaque monkey have been found that respond to motion stimuli containing elements of expansion, contraction, and rotation (Graziano, Andersen, & Snowden, 1994; Sakata, Shibutani, Kawano, & Harrington, 1985; Sakata, Shibutani, Ito, & Tsurugai 1986; Saito, Yukie, Tanaka, Hikosaka, Fukada, & Iwai, 1986; Tanaka, Hikosaka, Saito, Yukie, Fukada, & Iwai (1986); Tanaka, Fukada, & Saito (1989); Tanaka & Saito, 1989). There is also strong psychophysical evidence for a system in primate cortex processing these types of complex motion patterns (Freeman & Harris, 1992; Regan, 1986; Spitz, Stiles, & Siegel, 1993). It has been proposed that such pattern selectivity may be important in the tasks of ego-motion representation and the analysis of object motion in the environment.

MSTd is thought to be part of the motion processing stream that courses dorsally in cortex from V1 to MT to area MST (Boussaoud, Ungerleider, Desimone, 1990). V1 and MT both contain units tuned to linear motion (Albright 1984; Hubel & Wiesel, 1962; Livingstone & Hubel, 1988; Maunsell & Van Essen, 1983a,b) and the selectivity of MSTd cells to more complex motion patterns is thought to be built up from these more simple inputs. It is likely that motion direction and speed discrimination are processed together in the same cortical pathway. Recent studies (Pasternak & Merigan, 1994) have showed that lesions to the fundus of the superior temporal sulcus (STS) known to effect both areas MT and MST have raised both speed and motion direction detection thresholds for noisy stimuli.

The distribution of units in MSTd tuned to different motion patterns is biased in favor of expansion. Many more cells are tuned to stimuli containing expansion than either clockwise or counter-clockwise rotation, by a ratio of about 2:1 (Duffy & Wurtz, 1991; Graziano, *et al.* 1994; Saito, *et al.* 1986; Tanaka & Saito, 1989). The one exception to this general observation was reported by Lagae, Raiguel, Xiao, and Orban (1994), where more cells tuned to rotation than expansion were found. Given the consensus established by the other groups, we agree with their suggestion that the discrepancy is a consequence of the anatomical clustering of neurons with similar response selectivies in MSTd. Because this latter study involved relatively few penetrations, its population estimates were particularly vulnerable to sampling bias.

Given this strong anisotropy in response selectivity, it would be reasonable to expect perceptual consequences. A priori, lower discrimination thresholds for either speed or direction discrimination might be anticipated for motion stimuli containing elements of expansion. However, Sekuler (1992) found little evidence for this and concluded that speed detection thresholds were similar for linear motion, expansion, and contraction and could be predicted on the basis of sensitivities to the local motion signals that made up these stimuli.

In this study we take a different approach and compare perceived speed differences between stimuli containing expanding or rotating global motion signals. (This was originally motivated by the informal observation during a MSTd single unit recoding study that the random dots present in expansion patterns appeared to have a greater speed than those in rotation patterns even though the patterns had identical speed distributions.) Although the following psychophysical study uses human subjects while the physiology studies showing a preponderance of expansion units were done in the macaque, we believe that the cross-species generalizations have legitimacy. Whether MSTd has a role in ego-motion or object motion representation, expansion is arguably a more biologically relevant motion pattern than either contraction or rotation and it is not surprising that these units are over-represented. We also believe that this perceptual phenomenon is of interest, independent of this specific explanation.

GENERAL METHODS

The following conditions were adhered to unless otherwise specified for a particular experiment. Stimuli were generated and data collected on a

Macintosh computer with a 13 inch color Trinitron monitor. In all cases subjects viewed the stimuli 24 inches (61 cm) away from the display in a moderately lit room. In many cases the size of the stimulus was a circle of diameter 200 pixels. At the viewing distance used, this corresponded to a viewing angle of 7.63 degrees. The random-dots comprising each stimulus were plotted into a virtual square with dimensions 200x200 pixels. To obtain a circular stimulus presented on the screen, a circular mask was used to limit those pixels visible. A square stimulus was considered undesirable because of transiency problems evident in rotation patterns as dots rotate off the edge of the stimuli. Using a circular mask eliminated this problem for the case of rotation. With the expansion stimuli, dots could still move off the boundary of the visible display before living out their full life-times, a problem which is discussed and controlled for below under Experiment 6. Other masks were employed, in particular, in Experiment 5, where double wedges of various angular extents were applied.

Each dot was a square pixel that extended over a visual angle of 0.038 degrees per edge. Unless otherwise indicated, each pixel that was "on" to define a dot in the stimulus was a small black square against a white background. This arrangement eliminated persistence artifacts associated with bright moving features over a dark background. The fixation point was a filled circle of diameter 5 pixels (0.2 degree at the viewing distance). The refresh rate of the video card was 60.0 Hz. Each refresh cycle generated a software interrupt signal that caused the animated sequence of the stimulus, or "movie", to advance by one frame. Accordingly, a one second movie consisted of 60 consecutive image frames. To conserve memory, if the stimulus lasted longer than one second, it was started over from frame 1, in a loop. In each stimulus, the life-time of the dots was constrained to a limited number of frames (generally 12). Once a dot had persisted for this period of time, it was randomly assigned a new starting position, trajectory, and speed, consistent with the global motion pattern of the stimulus (see below). For the first frame of a movie, a random age was assigned to each dot, ranging from 0 to one frame short of being extinguished. This caused dots to "die" asynchronously and prevented a global flicker of the pattern with a periodicity determined by the life-time of the dots. If the dot left the virtual square defining the stimulus boundary, it was given a new random location within the stimulus, whether or not it had completed its entire life cycle. This prevented any fluctuation in dot density across the pattern from frame to frame. Note that premature dot “death” does not necessarily correspond to a dot moving outside the visible portion of the stimulus. It was generally the case that only a portion of the virtual square pattern was visible, as a result of applying an occluding "window" or mask over the movie.

For each paradigm, except that of Experiment 7 where this variable was specifically manipulated, dot density was the same for all the experiments. 100 dots in the 200x200 virtual stimulus square were shown at any one point. This meant that 1 out of 400 pixels were "on" at a time. Because various masks were used, not all of these pixels were visible. For example, when a 200 pixel diameter circular window was applied, 78.5% (31416 out of a possible 40000 pixels) were visible to the observer. Under this condition, an average of ~78 dots were visible each frame.

In most of the stimuli, the speed of each dot was proportional to the distance of its starting point from the center of the pattern. In terms of the motion of individual dots, there was no component of local acceleration in the signal. Not only did individual dots not deviate from their linear trajectories during a life-time, the speed of the dot was also constant. This is, of course, inconsistent with the movement of features on real objects expanding and rotating in the world. The "no acceleration" constraint was imposed to allow precise matching of local velocity vectors between stimuli with different motion patterns. For example, to go from a random dot display with local motion vectors organized into a global expansion (i.e. all velocity vectors pointed away from the center of the stimulus) to global rotation, all that needs to be done is to rotate each of these local vectors by 90 degrees (Graziano, *et al.* 1994). Figure 1 shows why this is the case. Clearly the average speed of the dots for the patterns in "A" and "B" are perfectly matched -- to transform one pattern into the other, a rotation of 90 degrees of each local motion vector is all that is required. Furthermore, on average, the same distribution of velocity vectors is present in the two patterns, only the spatial organization of these vectors has changed, allowing for a precise balancing of signals. Although both curvature and acceleration were absent from the local motion of the individual dots, this was not noticeable because the life-time of each dot was kept brief. If dot life-times become too long this no longer holds. For example, in the case of rotation, without curvature the dots would move progressively away from the center of the stimulus, giving the perception of an expanding spiral. The same problem would occur if dot speed were allowed to increase, as the phenomenon is directly dependent on the amount of spatial displacement. Dot life-times and speeds in this study were well beneath the point where this effect becomes noticeable.

A two alternative forced choice (2AFC) paradigm was used in all of our experiments. Subjects initiated each trial by pressing the "space" bar on a computer keyboard. They were told not to press this key until they were looking at the fixation point. Eye position was not monitored, as the effect seems largely independent of this behavioral variable. Head position was also not fixed, allowing for greater comfort for the subject. Although we insisted that the participants maintain a fixed viewing distance (24 inches), this variable in pilot studies was shown to have little effect on the outcome of the experiments. Following trial initiation, the first stimulus appeared at the center of the display. The center of the stimulus and the fixation point were transparently superimposed. This first movie was followed by a one second gap, where only the fixation point remained on the screen. This gap was followed by the second stimulus, which was presented in the same manner as the first. After the presentation of the second stimulus, both fixation point and movie were extinguished. At this point, the subject had to decide which stimulus had dots moving at the greater average speed. Participants were urged to ignore all aspects of the stimulus except the average speed of the whole patterns. They were also discouraged from formulating their judgments on the movement of individual dots. Based on this comparison of average speed, either '1' or '2' was pressed on the keyboard to register the choice, depending on whether the first or second stimulus had dots with greater perceived speed. Subjects were encouraged to take breaks from the task if they felt themselves becoming fatigued. Generally a few training trials were allowed prior to data collection. At no time was any feedback given to the subject about performance. For Experiments 2-7, four or five subjects were used, 2-3 naive observers as well as the two authors. For the first experiment, 3 additional naive participants were employed.

For each experiment, a "standard" expansion stimulus appeared as one of the two movies for comparison in each trial. The other movie in a trial was chosen from a set of "test" rotation patterns with dot speeds equal to 70, 80, 90, 100, 110, 120, and 130 percent of those present in the standard movie. In half of the trials the standard movie appeared first while in the remaining half one of the test patterns appeared first. All trials in an experiment were presented in a psudorandom order. By plotting the frequency with which each rotation pattern appeared moving faster than the standard expansion pattern as a function of the actual rotation to expansion speed ratio, an equivalence point could be recovered by fitting the data to a logit function and obtaining the 50% judgment point. In a few cases, this point lied outside the range of speeds tested and an extrapolation was required to obtain an assessment of the perceptual equivalence point.

Because of the 2AFC design of the experiment, exact 95% confidence intervals for the data points could not be established. It is not possible to produce a binomial confidence interval that will satisfy the strict definition of a confidence interval, namely one that will have the specified probability P of containing the unknown but fixed parameter p. For example, if the actual probability of a particular event is zero, any confidence interval established for such a point must necessarily include zero within its bounds, making the concept of a “95%” or “99%” confidence interval senseless. This problem arises because the observed probabilities for each data point can only take on discrete values. Although probability estimates for binomial data don't follow a normal distribution, they approach this form for large N, and by using a "continuity correction" confidence intervals were estimated by the standard methods (Snedecor & Cochran, 1989).

The broken curves bracketing the solid regression lines in Figures 4, 6, 8, and 9 represent the 95% confidence bands for the data. They were obtained by fitting a curve (using the logit function) to the upper and lower bounds of the 95% confidence intervals, respectively, associated with the data points. The dotted drop-lines extending downward from these curves bracket the solid drop-line associated with the function (also a logit function) obtained from the "best guess" data points. This technique will be used throughout this paper to get an estimate of the uncertainty associated with measuring each equivalence point. An effect will be considered "significant" if this interval does not include the "no effect" condition.

Experiment 1

Rational/Methods

Two basic stimulus patterns were used in this paradigm, which demonstrated the basic finding of this investigation. The stimuli used are shown in Figure 1. In this case, a 200 pixel circular mask was used to restrict viewing to a round window. The standard movie was an expanding dot pattern with dot life-times of 12 frames, or 200 msec. Dot speed for a particular pixel in the standard stimulus was established according to the formula: speed = k*(distance from the origin). In all cases, k was fixed at 0.02, which meant that the speed of a dot at the very edge of the stimulus window was 2 pixels/frame or 4.6 deg/sec. In the unlikely event that a dot happened to appear in exactly the center of the display, its velocity would be zero throughout its life-time. The velocity field which resulted is equivalent to that of an approaching circular object with a flat surface normal to the line of sight. However, because the velocity field did not change over time, the simulated distance of this object remained unchanged, as did the size of the stimulus, i.e. the stimulus did not evolve. As discussed above in General Methods, by rotating each velocity vector defining the expansion by 90 degrees to the left, a counter clockwise rotating pattern was achieved without altering the average speed distribution. These rotation patterns had an angular velocity of 68.7 degrees/sec. Rotation stimuli of various average speeds, both slower and faster, were also created to complete the set of "test" patterns, as discussed above. It should be pointed out that when we refer to a distribution of velocity vectors as being "identical" we mean statistically identical and not literally so. Because every dot for each pattern is randomly assigned a location, we do not literally rotate the exact same set of vectors in transforming one stimulus pattern into another. However, because the number of these random events is large in constructing these stimuli, we were not concerned that stochastic fluctuations in the average speed represented would have any effect on the results.

Results

Figure 2 shows what the performance curves would look like if there was no effect of motion pattern on perceived dot speed. The fraction where the rotation stimulus was judged "faster" is plotted against the ratio of Rotation Speed to Expansion Speed. The series of logit functions shown in the graph represent the performance of the same ideal observer if progressive amounts of noise was added into his perceptual system. The less noise added, the steeper the slope at the inflection point. A step function (not shown) which takes a jump from 0 to 1 at a speed ratio of 1.0 (physical equivalence) represents the case of ideal subject performance with no added noise. The horizontal line extending from the vertical axis bisects this axis at the 50% judgment point. By following this line over to the performance curve and then down to the abscissa, the point of perceptual equivalence can be recovered. For the "no effect" case depicted in this figure, this is of course a speed ratio of 1.

Figure 3 shows the actual experimental results. Each frame in this figure shows data collected from a single subject. The authors served as two of the subjects, the remaining six were unpaid volunteers naive as to the purpose of the study. To varying extents, each subject's psychophysical curve is shifted rightward from the “no effect” case. In each graph, the perceptual equivalence point (shown in the lower right corner of each frame) was greater than unity, indicating that all eight subjects perceived the dots in the expansion pattern moving faster than those in the rotation pattern. The bars drawn for each data point represent 95% confidence intervals. Figure 4 shows the data from the eight subjects used in Figure 3 pooled into a single curve. From this last plot, it is seen that the equivalence point for the set of subjects as a whole was a speed ratio of 1.21. In other words, the dot speed for the rotation pattern needed to be increased 21% before the perceived speed was the same as for the expansion pattern.

Experiment 2

Rational/Methods

We decided to systematically explore which aspects of the stimuli were responsible for the speed illusion documented in the first experiment. There were at least two components to the global organization of the velocity vectors defining the previous patterns. The original patterns had both a direction and speed gradient. In these movies, the speed of each dot was a linear function of its distance from the center of the display. In this second experiment, we eliminated this aspect of the stimulus, giving each dot a fixed speed, regardless of location. Two representative velocity fields from these patterns are shown in Figure 5. The speed of each dot was the same as a dot in the patterns from Experiment 1 located 71 pixels away from the center of the display. In this way, the average speed of the dots in the two experiments was approximately the same, although this was a relatively unimportant detail since these different types of patterns were not directly compared. We call these new stimuli "direction fields" to distinguish them from the "velocity fields" explored previously. This terminology was used in an MSTd physiology study by Tanaka, *et al.* (1989). This group probed the response profiles of MSTd units to both direction fields and velocity fields and achieved nearly equally vigorous responses.

Results

Figure 6 shows the data organized into the same plot format as the previous experiment. For brevity, although discrete data points from all four subjects are plotted, the curve from this figure was obtained by pooling data across all subjects. The speed ratio equivalence points for individual subjects were bg = 1.14, nq = 1.18, eg = 1.28, yz = 1.18. As seen from the plot of the pooled data, the overall equivalence point was a speed ratio of 1.19 and this effect was significant. In this case, dot speeds for the rotation pattern had to be, on average, 19% faster than the dot speeds for the expansion stimuli to appear to move at the same speed. Although the illusion was slightly less for the direction field compared to the velocity field in Experiment 1 (1.19 compared with 1.21, with overlapping confidence intervals), in each case the curves deviated significantly from veridical expectations. We concluded that the speed gradient contributed relatively little to the illusion, which is consistent with MSTd cells responding only slightly less to patterns where this component had been removed (Tanaka, *et al.*, 1989).

Experiment 3

Rational/Methods

Tanaka, *et al.* (1989) showed that the responses of MSTd cells to complex motion pattern trailed off as the number of local motion directions defining these patterns decreased. In the extreme case, that of axial expansion and rotation, only two directions of motion are present and responses evoked are quite weak. We predicted that if activity in MSTd was accounting for the illusion documented above, the strength of this effect should be smaller for stimuli which largely fail to activate this region. The patterns used are diagrammed in Figure 7. Frame A shows an example of axial expansion and rotation used for this study. Within a stimulus, all dot speeds are equal and only two directions of motion are represented in each pattern. As in the other experiments, the expansion stimulus was used as the standard in the 2AFC task. To transform the expansion pattern into a rotation pattern with identical velocity distributions (but not identical spatial organization) the expansion stimulus was effectively bisected orthogonal to its long axis. The left half of the stimulus was then placed on top of the right half, creating the axial "rotation."

Because of the way these patterns were constructed, the expansion stimulus was 100 pixels wide and 100 pixels high while the rotation pattern was 200 pixels wide and 50 pixels high. Because the shape of the two pattern types was not identical and their motion borders differed in length, an unwanted variable was introduced that could potentially effect the perception. To control for this, we created axial expansion and rotation patterns like those in Frame B. In these stimuli, the expansion patterns were oriented horizontally and the rotation patterns were square. By pooling the data from these two stimulus sets, the confounding effect of stimulus dimension was eliminated (this assumes there is no interaction between the two possible effects.)

Results

Unlike the stimuli in the first two experiments, no consistent effect of motion pattern on perceived average speed was evident. Figure 8A shows the individual data and a pooled tuning curve for the five subjects where stimuli like those of Figure 7A are compared. The individual subjective equivalence points were bg = 1.09, dk = 1.01, nq = 1.12, eg = 0.84, yz = 1.12. The equivalence point obtained from pooling this data was 1.03, a much smaller effect than that reported above for the isotropic patterns. Furthermore, the range of equivalence points bracketed by the 95% confidence bands includes the “no effect” case. Figure 8B shows pooled data obtained from trials comparing the stimuli of Figure 7B. The individual subjective equivalence points were bg = 0.98, dk = 0.89, nq = 1.04, eg = 0.84, yz = 0.89. The pooled equivalence point in this condition was 0.94 and again the “no effect” case was included within the range bracketed by the confidence bands. Figure 8C plots the result of pooling data shown in frames A and B. The individual subjective equivalence points in this final case were bg = 1.04, dk = 0.95, nq = 1.08, eg = 0.84, yz = 1.00. The equivalence point obtained by averaging over all five subjects was 0.99, indicating that the speed illusion previously obtained for isotropic patterns was not present for axial patterns. Consistent with implicating units in MSTd with the "speed illusion," axial expansion/rotation stimuli, which poorly activate this region, failed to elicit the speed illusion. We conclude that the presence of a wide range of directions in the original patterns used in Experiment 1 is responsible for the speed illusion. Note that this experiment also suggests that centrifugal organization (away from the line of sight) of motion vectors, per se, does not contribute to the phenomenon. Despite possessing more centrifugally oriented local motion signals in the axial expansion displays compared to rotation, the perceived dot speed was the same. We examine this issue further in the next experiment.

Experiment 4

Rational/Methods

Two competing hypotheses could explain the data obtained from the first two experiments. One possibility, discussed in the Introduction, suggests that the disproportionate number of pattern motion detectors in MSTd tuned for expansion compared to rotation is the basis for this illusion. Alternatively, a similar anisotropy with regards to local motion detectors could equally explain the results. Albright (1989) has reported that neurons in area MT tuned for local linear motion oriented away from the fovea (centrifugally) are over-represented. Because of this, we would expect that the expansion patterns would activate more units in MT than stimuli with any other distribution of local motion signals. If this were the case, an anisotropy with regards to complex motion pattern detectors in MSTd need not be invoked to account for the illusion.

To distinguish between these two alternatives, we altered the experimental paradigm so that rather than presenting the two stimuli being compared sequentially, they were presented together, side by side, with an intervening gap of 48 pixels (1.83 degrees). In the center of this gap the fixation point was presented. By placing the stimuli in the periphery, on average the patterns all had the same number of centrifugally oriented local motion vectors. Consequently, the expansion and rotation patterns positioned at these eccentric locations should equally activate the population of MT cells. Because this task was much more difficult, due to the eccentrically placed stimuli, we showed the movies for a full 3 seconds (actually showing the same one-second movie repeated 3 times without an intervening gap). Subjects pressed "1" or "2" depending on whether the movie to the left or right of fixation, respectively, appeared to have faster moving dots.

Results

Figure 9 shows the results for this experiment, using data pooled over four subjects. The individual subjective equivalence points were bg = 1.20, nq = 1.30, eg = 1.39, yz = 1.23. The overall effect was somewhat larger than that observed in the first two studies, with the equivalence point from the pooled data established at 1.27, significantly above the “no effect” condition. We conclude that the anisotropy in area MT cells favoring centrifugally oriented linear motion detectors could not explain the illusion.

Experiment 5

Methods/Rational

Based on the results of comparing axial patterns in Experiment 3, we predicted that by extending the range of local motion directions which defined the motion patterns, we should expect to see an increasing rightward shift in the perceptual equivalence speed ratios. To test this hypothesis, we constructed double "wedge" patterns as shown in Figure 10. We used the same rules established for the stimuli in Experiment 1 (the "velocity fields") for the movement of the random dots, but instead of applying a circular mask over the square virtual dot patterns, two wedge shaped masks were used instead. Wedge pairs were centered around the vertical line through the center of the display and the two wedges were oriented 180 degrees away from each other, with the rounded edges of the wedges oriented up and down. The 2AFC task was run with the expansion pattern as the standard compared against various rotation patterns with different dot speeds. This was repeated for wedges of angles 30, 60, 90, 120, 150, and 180 degrees. A wedge pair of 180 degrees is equivalent to two semi-circles and therefore was identical to the circular patterns of Experiment 1. For this data point, we used the previously collected data rather than repeat the identical study. Only wedges of the same size were compared with one another.

Results:

Figure 11A shows the pooled data collected from five subjects. The six curves were obtained by regressing data collected from showing each of the six wedge sizes separately. Rather than show the entire curve, a small portion of the X axis has been expanded to show the shift of equivalence points more clearly. Figure 11B shows this data organized into a different format. In this plot, subjective equivalence ratio is plotted as a function of stimulus wedge size. A clear trend is evident in both these graphs: the larger the area of the stimulus exposed, the more a subject's judgment of greater speed magnitude favored the expansion pattern. Two-way ANOVA was performed on the entire data set, and a significant effect of wedge size on the equivalence point was established ($p<0.05$).

For small wedge sizes, a reversal of the illusion was seen for some subjects -- the rotation patterns were more frequently judged as possessing greater average speeds. We attribute this as arising from the phenomenon of "temporal capture" (Treue, Snowden, & Andersen, 1993). Dots in a rotation pattern will have, on average, shorter life-times because they may disappear prematurely off the two long side edges of the stimulus. In comparison, dots in the expansion pattern may only disappear prematurely at the much shorter top and bottom edges of the stimulus. This effect is reduced as wedge width increases and the effect of motion pattern on perceived speed quickly dominates. Although the trend in the data reflects two competing effects, the results are consistent with an increasing effect of global motion pattern with an increase in the distribution of motion vectors that define these patterns. This is also consistent with data from Tanaka, *et al* (1989) which demonstrated an increase in response strength for MSTd neurons with increasing number of local motion signals that define the stimulus patterns used to probe these units.

Experiment 6

Rational/Methods

We were concerned, given the results of the previous study, that we had not adequately controlled for dot transiency in our earlier experiments. For example, in Experiment 1 where a circular mask was used, the average life-time of dots in the expansion patterns is slightly less than for the rotation patterns. This is because dots near the periphery of the stimulus can move off the edge of the expansion but not the rotation stimuli. Although this difference in average life-time was small, because shorter average dot life-time produces greater perceived average speed, we were concerned that this could account for at least part of the illusion. To see if this was the case, we repeated the original experiment using dots with very short life-times -- four frames. As dot life-time is reduced, so is the number of dots prematurely disappearing off the edge of the stimulus. Another option for dealing with this transiency problem would have been to restrict the starting locations for each dot so that their entire path remained within the stimulus window. Unfortunately, this introduces other artifacts, such as differences in dot density near the periphery of the stimuli.

We decided to extend this study to include patterns with dot life-times of 6, 8, 12, and 16 frames, to see if we could establish a consistent trend. If differences in dot life-time between the expansion and rotation displays were accounting for the speed illusion, we would expect to see this effect to become larger at longer dot life-times. Life-times much longer than 16 frames were rejected in order to avoid pattern distortions of the type discussed in the General Methods section. The 2AFC task was again employed, as usual using a single expansion standard pattern compared against rotation patterns of varying dot speeds. Only stimuli possessing the same dot life-times were compared.

Results

Figure 12A shows data from five subjects for each of the dot life-time conditions. A curve is fit to the pooled data for each life-time. Frame B takes the subjective equivalence ratios for each set of patterns and plots them as a function of dot life-time. To our surprise, the shorter the dot life-time, the greater the magnitude of the illusion, the opposite of what would be expected if temporal capture were accounting for the effect. Two-way ANOVA was performed on the entire data set, and a significant effect of dot life-time on the equivalence point was obtained ($p<0.05$). It is clear that whatever effect differential transiency is having on the perceived speed of dots near the periphery of these patterns, it is small compared with the overall effect of global motion. Not surprisingly, those sets of patterns with shorter life-times appeared to have dots moving considerably faster than those patterns with longer life-time. However, this was not important because the patterns being compared always had the same dot life-times. Why patterns of shorter dot life-time gave a stronger effect will be considered in the Discussion.

Experiment 7

Rational/Methods

We pursued the line of reasoning in Experiment 5 in a slightly different direction. Rather than varying the area of the stimulus exposed as was done in the previous study, we systematically ran a series of experiments with different dot densities. 2, 4, 8, 16, 32, 64, 128, and 256 dots were used with patterns that were otherwise identical to those employed in Experiment 1. At low dot densities, problems associated with stochastic fluctuations in average speed potentially became an issue. To avoid this problem, the random number seed for the program generating the stimulus patterns was saved and reused before each movie was created. As a consequence, the initial spatial location of the random dots was identical for the patterns being compared.

Based on results from Experiment 5 which suggested a positive relationship between the range and number of local motion directions present in the patterns and the magnitude of the illusion, we predicted the difference in perceived speed would increase with the number of dots in the display. Because the dots were repositioned every 12 frames, the number of motion directions represented in the stimulus patterns over the duration of the movie was greater than the number of dots present at any one time on the screen. For example, in the 2-dot condition approximately 2*(60 frames)/(12 frames), or 10, different motion directions were sampled over the course of a 1 second movie.

Results

Figure 13 shows that the results were consistent with expectations. The magnitude of the illusion is considerably less for the two and four dot conditions than for the remaining cases. Frame A of this figure shows a series of eight curves, one for each dot density, for data pooled over four subjects. Frame B plots individual equivalence points for each subject along with a line connecting data points obtained from pooling this data. A clear positive correlation between dot density and the magnitude of the illusion is evident in this plot . Two-way ANOVA was performed, and a significant effect of dot density on the subjective equivalence point was established ($p<0.05$). The results are consistent with the rest of the data collected in this study: the illusion is directly correlated with the strength of the global motion pattern present in the stimulus.

DISCUSSION

In the Introduction we suggested that the reason local features defining expansion motion appear to move with greater speed than features defining rotation motion was because of the over-representation of expansion detectors in MSTd. We predicted that the magnitude of the perceptual effect would be well correlated with MSTd selectivity. In each experiment we performed, the results were consistent with expectations. Reducing the number of local motion directions defining expansion and rotation in Experiment 3 (down to two directions in the case of axial expansion/rotation) deminished the illusion, consistent with the poor responses reported when these patterns were used to drive MSTd units. Removing the speed gradients from the patterns, thus reducing them to "direction fields", had little effect on the speed illusion, consistent with the strong responses these patterns produce in MSTd neurons. The centrifugal bias of MT direction selectivity was effectively eliminated as a potential explanation for the illusion, as demonstrated in Experiment 4, where moving the patterns away from the fovea actually slightly increased the magnitude of the subjective speed difference between motion patterns. Additionally, the axial expansion patterns used in Experiment 3 contained a disproportionate amount of local motion signal with a centrifugal orientation compared with the axial rotation patterns and no effect was reported. These axial patterns also do not activate MSTd neurons well. We conclude that the speed illusion reported here is closely related to MSTd activity and that it is generated by the global arrangement of a wide range of motion directions.

We also found that reducing dot density in the stimuli had the effect of reducing the magnitude of the illusion. Unfortunately, there is no good corresponding physiological data on this issue. With regards to the effect of dot density on MSTd unit response profiles, lower dot densities than those used in this study still strongly stimulated cells (Duffy & Wurtz, 1991). However, because the stimuli they used were 100 degrees in diameter, the number of dots in the displays was never less than 25 at one time. It is not clear how these same units would have responded to the much smaller stimuli used in this investigation with many fewer dots.

Does the Primate Visual System Possess Complex Motion Pattern Detectors?

An assumption of this study is that the primate cortex possesses mechanisms specifically sensitive to complex motion pattern. This study further provides evidence for such a mechanism by showing an effect of complex motion pattern on perceived speed. As discussed above, Experiments 3 and 4, in particular, make it difficult to attribute the speed illusion to the functioning of a system of local linear detectors. Along similar lines, Graziano et al (1994) proposed that MST cells are similar to IT cells in being selective for patterns, but in the case of MST the selectivity is for patterns of motion rather than static patterns. Both areas demonstrate positional invariance with respect to pattern selectivity — the stimulus specificity is relatively insensitive to stimulus placement within a unit's receptive field. The receptive fields in both areas are very large and include the fovea, indicating that single cells generalize for patterns over a large extent of the visual field. And finally, both areas use inputs from areas analysing visual information at a much more local, rudamentry level. Thus, just as area IT has detectors for faces, area MST has detectors for expansions, rotations, and spirals.

The possible existence of a specialized motion pattern network remains far from resolved in the psychophysical literature. Two groups (Braddick & Holliday, 1991; Werkhoven & Koenderink, 1991) observed that the motion patterns of expansion, contraction, rotation, and shear do not "pop out" in displays containing distracters. One possible interpretation of this data is that the human visual system lacks a specialized motion pattern system. Instead, a selective attention mechanism is used to integrate patterns distributed across local, unidirectional detectors (such as those found in MT). However, pop-out results can be interpreted in a number of ways. For example, the large receptive fields and lack of a retinotopic organization to area MSTd is consistent with a system that, although specialized for analyzing complex motion patterns, cannot pre-attentively process multiple motion patterns that are spatially separated.

Experiments looking at speed discrimination thresholds for complex motion patterns have reinforced the possibility that the primate visual system lacks a specialized network for detecting these stimuli. These studies have shown that the thresholds for looming, rotation, and linear motion are all similar (Sekuler, 1992), providing evidence against separate processing channels for these different motion types. Consistent with these patterns having a distributed representation at the level of local detectors, thresholds for complex motion patterns were what would have been predicted based on the simple pooling of local, linear motion signals. Finally, a recent illusion reported by Duffy & Wurtz (1993), where a translational velocity field superimposed over an expansion pattern shifts the perceived focus of expansion in the direction of translation, argues for a lack of separation between channels that process expansion and linear motion.

On the other hand, a significant number of studies support the possibility of a complex motion network. Data from adaptation experiments (Regan, 1986) suggests the presence of independent channels tuned to linear motion, expansion, and rotation. Regan was able to develop stimuli which selectively increased perception thresholds for one pattern type without effecting the others. Consistent with a "low-level" processing of complex motion pattern, studies in infants (Spitz, *et al.*, 1993) have showed that the capacity to integrate the information contained within nonuniform velocity fields into coherent motion patterns develops as early as 7 months of age. Masking studies (Freeman & Harris, 1992) indicated that the detection of expansion in a stimulus is unaffected by the presence of rotation, suggesting at least two independent channels for processing complex motion patterns. However, these studies are difficult to reconcile with data from a physiology study (Orban, Lagae, Verri, Raiguel, Xiao, Maes, & Torre, 1992) which showed that such masking detunes MSTd cells.

Much of this controversy can be resolved by considering that many MSTd cells have multiple components to their selectivity. For example, many cells have been found that are tuned not just for expansion, but some direction of translational motion as well. Similarly, cells tuned for such combinations as contraction and clockwise rotation have been found as well. Although we believe that this region is involved in the representation of complex motion pattern, it is not surprising that a lack of independence has been found in these channels, given the mixing of signals known to occur at the level of individual units.

Since the speed illusion we reported in this paper cannot be explained by a local linear motion system, it provides further evidence for the existence of a complex motion pattern system.

Other Speed Illusions and Perceptual Anisotropies

Although more attention has been paid to the directionally selective mechanisms of velocity perception than to speed, the literature is scattered with reports of various speed illusions. Watamaniuk, Grzywacz, & Yuille (1993) noticed that increasing the density of features in translational random-dot stimuli led to increases in the apparent speed of the dots. This illusion was particularly strong when the increase in density occurred during continuous viewing of the stimulus, with a smaller effect reported once dot density was no longer changing. Along similar lines, Thompson (1982) reported that sine-wave gratings appear to move faster when they contain higher contrast. This same study found no dependency of speed perception on spatial frequency content but did find evidence that the orientation of the grating affected perceived speed. Another study also using drifting sinusoidal gratings found that these stimuli appear to move more slowly in the periphery than foveally (Johnston & Wright, 1986). As mentioned before, Treue, *et al.* (1993) reported that decreasing the dot lifetimes of stimulus features defining motion patterns causes an increase in the perceived velocity of these features. This effect was evident even when flickering dots which were not moving were interposed amongst the moving dots. Increasing the flicker rate of these stationary dots appears to increase the speed of the dots that are moving, even though the life-times of this moving group remain constant, a phenomenon which they call "temporal capture."

The only speed illusion we could find that involved rotating stimuli was in a report by Vicario & Bressan (1990) documenting the perception of rotating wheels on vehicles undergoing forward translation. They found that subjects consistently overestimate the angular velocity of the wheel relative to the forward velocity of the vehicle. This illusion creates the impression of the wheels partially "slipping" relative to the surfaces with which they are in contact. This is interesting, because given the results of this study, it might be expected that subjects underestimate the speed of rotating objects in general. It would be worthwhile to create random-dot displays of linear (translational) motion with identical speed distributions as the dots in rotation and expansion displays and compare estimated speed magnitudes.

There are numerous reports of perceptual distortions in the human motion processing system. Thresholds for the detection of coherent motion in displays with low signal-to-noise ratios are generally higher along the vertical meridian, particularly for motion moving either upward or downward (van de Grind, Koenderink, van Doorn, Milders, & Voerman, 1993). Another study (Raymond, 1994) reported that although foveal motion sensitivity was isotropic, a small but significant (about 0.1 log units) difference in sensitivity in favor of centripetal motion was observed at eccentricities between 5.0 and 12.5 degrees out from the fovea. This was true for the entire horizontal meridian and the inferior half of the vertical meridian. Motion sensitivities for the superior portion of the vertical meridian were isotropic (i.e. identical for all motion directions.) Consistent with the previous study, motion thresholds were generally higher along the vertical axis.

The phenomenon reported in the current study cannot be explained by any combination of the above factors, because the effect was invariant with regards to stimulus placement relative to the fovea. This is important because it shows a dependence of perceived speed on the global organization of the motion vectors that compose the stimulus. While detection thresholds and discrimination thresholds have traditionally been examined to probe the working of the nervous system, this investigation shows that systematic differences in perceived speed magnitude can also reveal information about the internal processing of the visual system.

Number of Cells and Perceived Speed

Although it seems plausible that a larger number of MSTd cells tuned to expansion compared to rotation would lead to expanding stimuli being perceived as moving faster than rotating ones, a real explanation of the illusion requires a computational model that relates MSTd cell activities to the perceived global pattern speeds. Unfortunately, such a model does not yet exist. Nevertheless, we would like to speculate on the types of mechanism that might lead to the illusion.

One possible explanation for the effect we reported is that differences in the proportion of units tuned to rotation and expansion could lead to different rates of fatigue for these two channels. Fewer units conceivably could leave a channel more susceptible to fatigue. In a pilot study, this explanation was ruled out by demonstrating that stimulus exposure time had little effect on the magnitude of the speed illusion. We tried exposure times of both one half and two seconds, bracketing the one second stimulus time used in this study.

Many models for local translational velocity computation have been proposed in the past (Horn & Schunck, 1981; Hildreth, 1984; Heeger, 1987; Grzywacz & Yuille, 1990; Simoncelli, 1993). These models cannot adequately predict our speed illusion because, as we have demonstrated, the illusion is a global phenomenon depending on the overall arrangements of many different directions of motion and it disappears when the global patterns we used are viewed through narrow, wedge-shaped apertures. However, if we assume that the computation of global pattern speed involves similar steps as in some physiologically-inspired local translational velocity models (Heeger, 1992; Heeger, 1993), our speed illusion could be explained. A key element in these motion models is a normalization step at which the output of a specific translational motion mechanism is divided by the sum of outputs of all the translational motion mechanisms. We could generalize this procedure to the case of global pattern speed computation by assuming that the output of the expansion (or rotation) mechanism is normalized by the outputs of all global motion mechanisms present in MSTd. It is also reasonable to assume that the signal strength of a given global motion mechanism before normalization is proportional to the number of MSTd cells tuned to that global motion type. Because there are more MSTd cells tuned to expansion than rotation, the output of the expansion mechanism after normalization would remain stronger than the rotation mechanism. This could be the physiological basis of the speed illusion reported in this paper.

A complication to the above explanation comes from the result of our Experiment 6, which indicates that the magnitude of the speed illusion increases under conditions of shorter dot life-time. To explain this effect, we assume that decreasing dot life-time disproportionately reduces local translational motion signals relative to global motion signals. This is because of the expected greater spatial and temporal integration capabilities of the latter system, which pools data over larger spatial and temporal windows. The observed effect of dot life-time on the speed illusion can be explained by interpreting perceived speed as the result of competition between a local and a global speed estimation process. At a local level, an estimate of speed is obtained for each dot based on its individual motion. This signal is based only on the local trajectory of the feature and is independent of the global arrangement of velocity vectors. V1 and MT are likely sites for such a local speed estimation. The global speed estimation may be assumed to be performed in MSTd where the overall arrangement of many local velocity vectors is taken into account. We do not know the nature of the operation that combines the outputs of the two regions but the known two-way connections between MSTd and MT might provide the necessary circuitry. The illusion's dependence on dot life-time suggests that under environmentally relevant conditions, where features are not so transient, the effect would be much reduced. We could not test stimuli with dot life-times extending beyond 0.25 seconds (16 frames) because of distortions, discussed in General Methods, introduced into our patterns.

A biological function of the proposed interaction between local and global motion estimation processes could be a gradual refinement of both processes. Knowledge of the global motion patterns could refine local motion estimation by providing better motion constraints than the commonly used uniform translational motion assumption. Improved local motion estimation could in turn refine global motion computation which receives locally computed velocity fields as input. It would be interesting to test this idea in an integrated model of simultaneous local and global motion computation.

Acknowledgments

We would like to thank Professor Richard A. Andersen for his invaluable support and his comments on an earlier version of this manuscript. We would also like to thank the naive subjects who agreed to serve as observers, particularly Ellen Geesaman, David Kane, and Yudong Zhu for partipating in the complete set of studies. Ning Qian. is supported in part by a research grant from the McDonnell-Pew Program in Cognitive Neuroscience. This work was also funded by NIH grant EY 07492 and the Human Frontiers Scientific Program awarded to Professor Richard A. Andersen.

REFERENCES

Albright, T. D. (1984). Direction and orientation selectivity of neurons in visual area MT of the macaque. Journal of Neurophysiology, 52, 1106-1130.

Albright, T. D. (1989). Centrifugal direction bias in the middle temporal visual area (MT) of the macaque. Visual Neuroscience, 2, 177-188.

Boussaoud, D., Ungerleider, L. G. & Desimone, R. (1990). Pathways for motion analysis: cortical connections of the medial superior temporal and fundus of the superior temporal visual areas in the macaque. Journal of Comparative Neurology, 296, 462-496.

Braddick, O. J. & Holliday I. E. (1991). Serial search for target defined by divergence or deformation of optic flow. Perception, 31, 345-354.

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 Neurophysiology 65, 1329-1345.

Duffy, C. J. & Wurtz, R. H. (1993). An illusory transformation of optic flow fields. Vision Research, 33, 1481-90.

Freeman, T. C. & Harris, H. G. (1992). Human sensitivity to expanding and rotating motion: Effects of complementary masking and directional structure. Vision Research, 32, 81-87.

Graziano, M. S. A., Andersen, R. A. & Snowden, R. J. (1994). Tuning of MST Neurons to Spiral Motions. Journal of Neuroscience, 14, 54-67.

Gryzwacz, N. M. & Yuille, A. L. (1990). A model for the estimate of local image velocity by cells in the visual cortex. Proceedings of the Royal Society of London B, 239, 129-161.

Heeger, D. J. (1987). Model for the extraction of image flow. Journal of the Optical Society of America A, 4, 1455-1471.

Heeger, D. J. (1992). Half-squaring in responses of cat striate cells. Visual Neuroscience, 9, 427-43.

Heeger, D. J. (1993) Modeling simple-cell direction selectivity with normalized, half-squared, linear operators, Journal of Neurophysiology, 70, 1885-1898.

Hildreth, E. C. (1984) Computations underlying the measurement of visual motion. Artificial Intelligence, 23, 309-355.

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

Hubel, D. B. & Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. Journal of Physiology (London), 160, 106-154.

Johnston, A. & Wright, M. J. (1986). Matching velocity in central and peripheral vision. Vision Research, 26, 1099-1109.

Lagae, L., Maes, H., Raiguel, S., Xiao, D. K. & Orban, G. A. (1994). Responses of macaque STS neurons to optic flow components: a comparison of areas MT and MST. Journal of Neurophysiology, 71, 1597-626.

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

Maunsell, J. & van Essen, D. C. (1983a). Functional properties of neurons in middle temporal visual area of the macaque monkey. I. Selectivity for stimulus direction, speed, and orientation. Journal of Neurophysiology, 49, 1127-1147.

Maunsell, J., van Essen, D. C. (1983b). Functional properties of neurons in middle temporal visual area of the macaque monkey. II. Binocular interactions and sensitivity to binocular disparity. Journal of Neurophysiology, 49, 1148-1167.

Orban, G. A., Lagae, L., Verri, A., Raiguel, S., Xiao, D., Maes, H. & Torre, V. (1992). First-Order analysis of optical flow in monkey brain. Proceedings of the National Academy of Science, 89, 2595-2599.

Pasternak, T. & Merigan, W. H. (1994). Motion perception following lesions of the superior temporal sulcus in the monkey. Cerebral Cortex, 4, 247-59.

Raymond, J. E. (1994). Directional anisotropy of motion sensitivity across the visual field. Vision Research, 34, 1029-37.

Regan, D. (1986). Visual processing of four kinds of relative motion. Vision Research, 26, 127-145.

Sakata, H., Shibutani, H., Kawano, K. & Harrington, T. L. (1985). Neural mechanisms of space vision in the parietal association cortex of the monkey. Vision Research, 25, 453-463.

Sakata, H., Shibutani, H., Ito, Y. & Tsurugai, K. (1986). Parietal cortical neurons responding to rotary movement of visual stimulus in space. Experimental Brain Research, 61, 658-663.

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.

Sekuler, A. B. (1992). Simple-pooling of unidirectional motion predicts speed discrimination for looming stimuli. Vision Research, 32, 2277-88.

Snedecor, G. W. & Cochran, W. G. (1989) Statistical methods. Iowa State University Press.

Simoncelli, E. P. (1993) Distributed analysis and representation of visual motion, PhD thesis.

Spitz R. V., Stiles J. & Siegel R. M. (1993). Infant use of relative motion as information for from: evidence for spatiotemporal integration of complex motion displays. Perceptual Psychophysics, 53, 190-9.

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., Fukada, Y. & Saito, H. (1989). Underlying mechanisms of the response specificity of expansion/contraction and rotation cells in the dorsal part of the medial superior temporal area of the macaque monkey. Journal of Neurophysiology, 42, 642-656.

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.

Thompson, P. (1982) Perceived rate of movements depends on contrast. Vision Research, 22, 377-380.

Treue, S., Snowden, R. J. & Andersen, R. A. (1993). The effect of transiency on perceived velocity of visual patterns: a case of "temporal capture". Vision Research, 33, 791-798.

van de Grind, W. A., Koenderink, J. J., van Doorn, A. J., Milders, M. V. & Voerman, H. (1993). Inhomogeneity and anisotropies for motion detection in the monocular visual field of human observers. Vision Research, 33, 1089-107.

Vicario, G. B. & Bressan, P. (1990). Wheels: a new illusion in the perception of rolling objects. Perception, 19, 57-61.

Watamaniuk, S. N. J., Grzywacz, N. M. & Yuille, A. L. (1993). Dependence of speed and direction perception on cinematogram dot density. Vision Research 33, 849-859.

Werkhoven, P., Koenderink, J. J. (1991). Visual processing of rotary motion. Perception and Psychophysics, 49, 73-82.

LEGENDS

Figure 1: Stimuli used in Experiment 1. Frame A shows an example of counter-clockwise rotation while Frame B is an example of expansion. Each arrow is a motion vector that represents the direction and magnitude of individual dots making up these patterns. Note that the length of these vectors increases moving outward from the center of the stimuli. As explained in the text, transforming one pattern into the other simply involves rotating each local motion vector by 90 degrees in the appropriate direction.

Figure 2: Series of ideal performance curves. The x-axis represents the actual speed ratios of a set of test rotation patterns to a fixed standard expansion pattern. The y-axis represents the fraction of trials in which a test rotation pattern is judged moving faster than the standard expansion pattern. If the two types of motion pattern being compared appear to move equally fast when their actual speeds are the same, the point of inflection of the logit function representing performance would be at a speed ratio of unity, as shown for the curves in the figure. The abscissal location of this point for real data shifts to the left or right depending on the subjective judgment of relative speed. The ordinal location of the inflection point is constrained by the general form of the logit function to be always at 0.5. The slope of the curve is inversely correlated with a particular subject's ability to consistently judge differences in speed.

Figure 3: Individual subjects' data from Experiment 1. Each plot shows the psychophysical performance curve for a different observer. In each case the point of perceptual equivalency is shifted to the right, indicating that each subject tended to judge dots in expansion patterns as moving more quickly. Error bars represent 95% confidence intervals.

Figure 4: Performance curve obtained from pooling data from Experiment 1 across subjects. broken curves are 95% confidence bands, as described in the text.

Figure 5: Stimuli compared in Experiment 2. These patterns are identical to those used in the previous paradigm, except that the radial speed gradient has been removed and speeds of all dots in a particular stimulus are identical.

Figure 6: Data collected using stimuli lacking a speed gradient. The rightward shift of the equivalence point is comparable to that obtained with patterns containing a speed gradient. Data from each of the four observers is plotted. This data was pooled for the purpose of obtaining the solid regression line. Broken flanking curves represent 95% confidence bands, as described in the text.

Figure 7: Stimuli used in Experiment 3. The patterns shown in Frames A and B represent examples of axial rotation and expansion, with only two directions of motion defining these global motion patterns. For a particular stimulus, the speed of all the dots was identical. The solid dot in the center of each pattern represents the fixation point.

Figure 8: Comparison of axial expansion and contraction. Frame A shows performance curves using stimuli like those in Figure 7A while Frame B uses

stimuli like those in Figure 7B. Frame C pools these previous two data sets. The difference in perceived speed between expansion and rotation disappears when axial patterns are compared.

Figure 9: Effect of moving stimulus patterns away from the fovea. The rightward shift of the inflection point is again consistent with expansion appearing faster. The effect was slightly larger than when the patterns were viewed foveally.

Figure 10: Stimuli used in Experiment 5. These patterns were created identically to those shown in Figure 1 except that a double wedge-shaped mask was applied.

Figure 11: Effect of using patterns of different wedge-size. Each curve in Frame A was obtained by pooling data across the five subjects tested. The six curves correspond to the six wedge sizes used. The 180 degree double wedge was identical to the stimuli used in Experiment 1 (a full circle). Frame B plots the subjective equivalence points for each subject as a function of wedge size. The solid line connects data points of the pooled data.

Figure 12: Effect of dot life-time on the magnitude of the illusion. Each curve in Frame A was obtained by pooling data across the five subjects tested. The five curves correspond to the different dot life-times. The 12 frame condition was identical to the stimuli used in Experiment 1. Frame B plots the subjective equivalence points for each subject as a function of dot life-time. The solid line connects data points of the pooled data set. Both ways of depicting the data show a clear relation between dot life-time and magnitude of the speed illusion, with the effect increasing at short life-times.

Figure 13: Effect of dot density on the location of the subjective equivalence point. Each curve in Frame A was obtained by pooling data across the four subjects tested. The four curves correspond to the different dot life-times. Frame B plots the subjective equivalence points for each subject as a function of wedge size. The solid line connecting points of the pooled data set shows a generally upward slope. Beyond the 4-dot case, this increase largely plateaus.

Figure 1

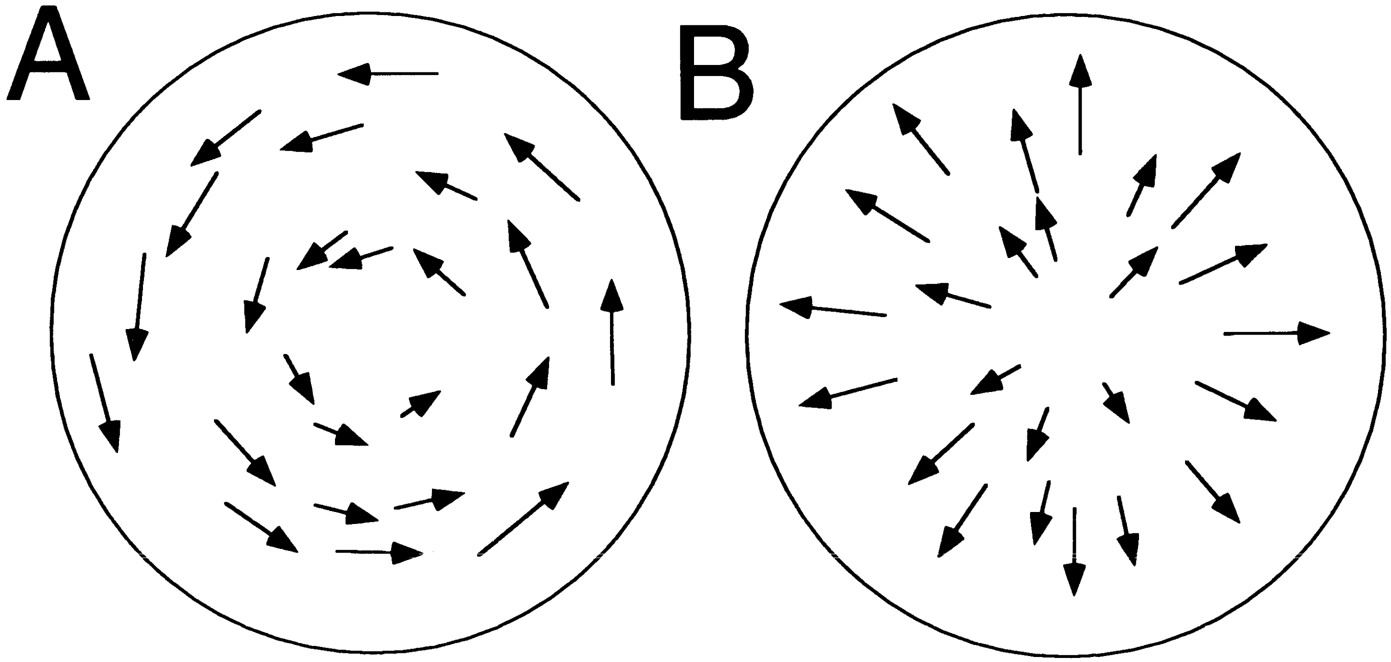
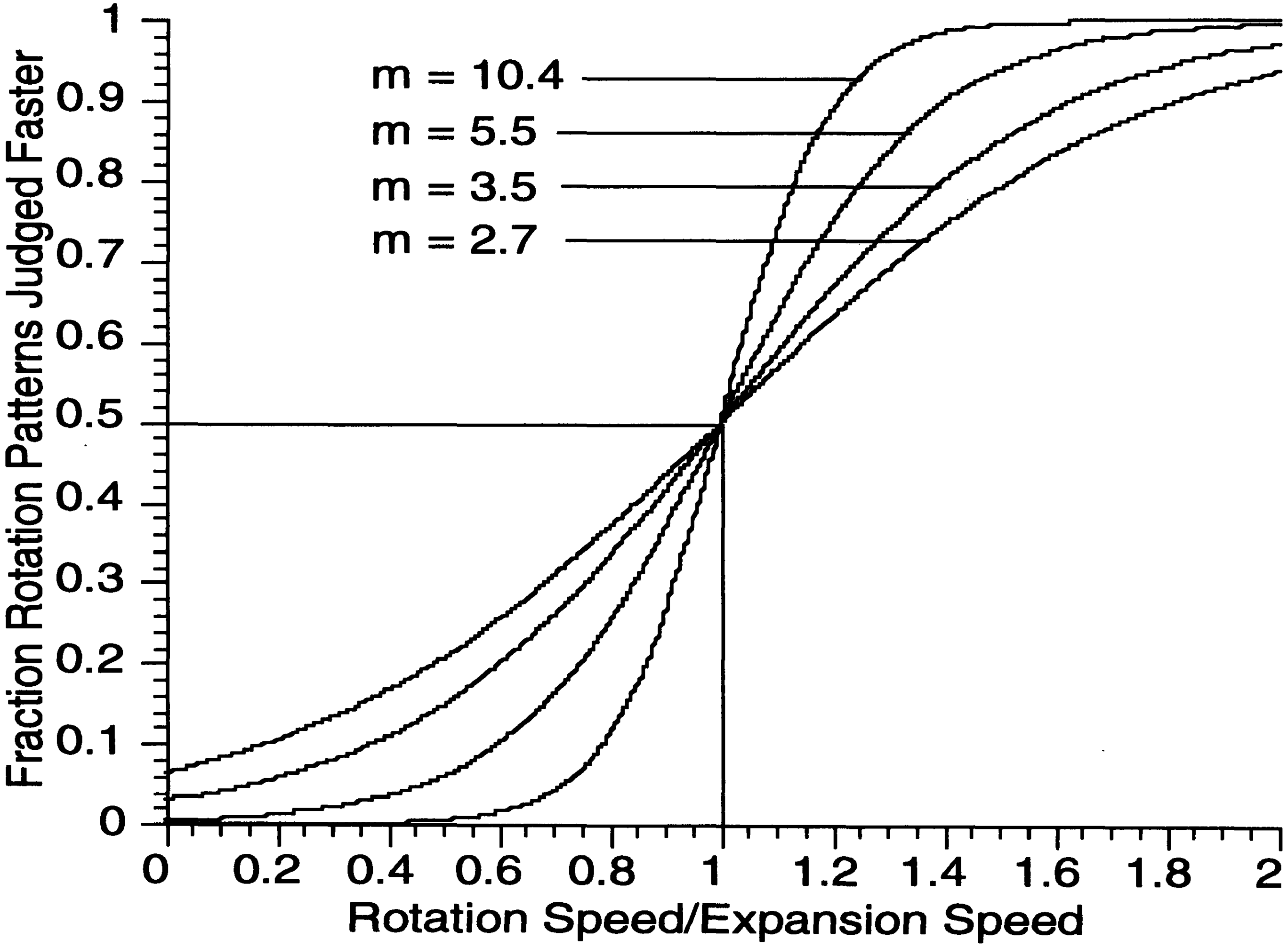
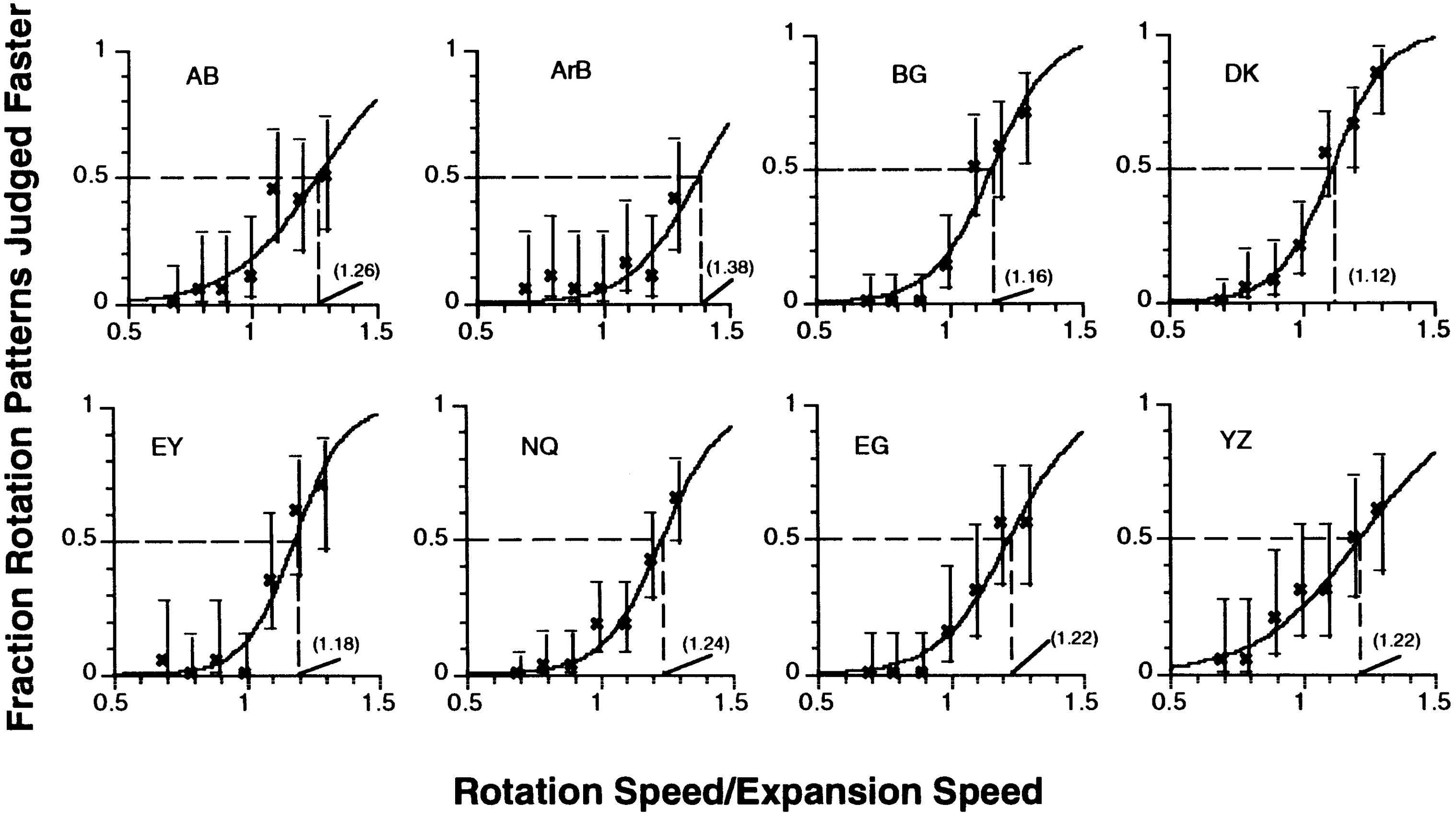


Figure 2



m = 10.4
m = 5.5
m = 3.5
m = 2.7
Fraction Rotation Patterns Judged Faster
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
0.2
0.4
0.6
0.8
1
1.2
1.4
1.6
1.8
2
Rotation Speed/Expansion Speed

Figure 3


Fraction Rotation Patterns Judged Faster
AB
(1.26)
ArB
(1.38)
BG
(1.16)
DK
(1.12)
EY
(1.18)
NQ
(1.24)
EG
(1.22)
YZ
(1.22)
0
0.5
1
0.5
1
1.5
Rotation Speed/Expansion Speed

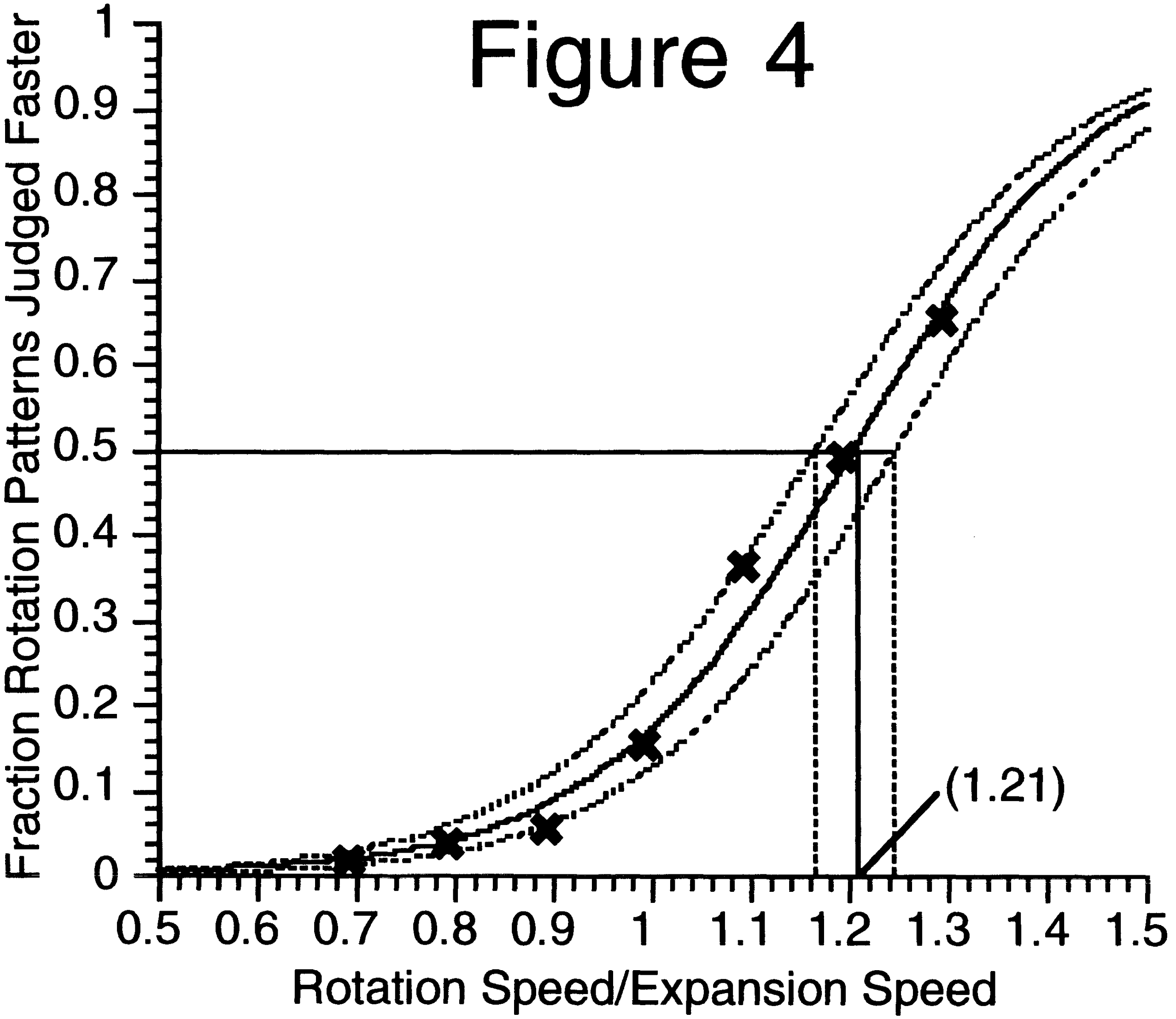


Figure 4

Fraction Rotation Patterns Judged Faster
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0.5
0.6
0.7
0.8
0.9
1
1.1
1.2
1.3
1.4
1.5
(1.21)
Rotation Speed/Expansion Speed

Figure 5

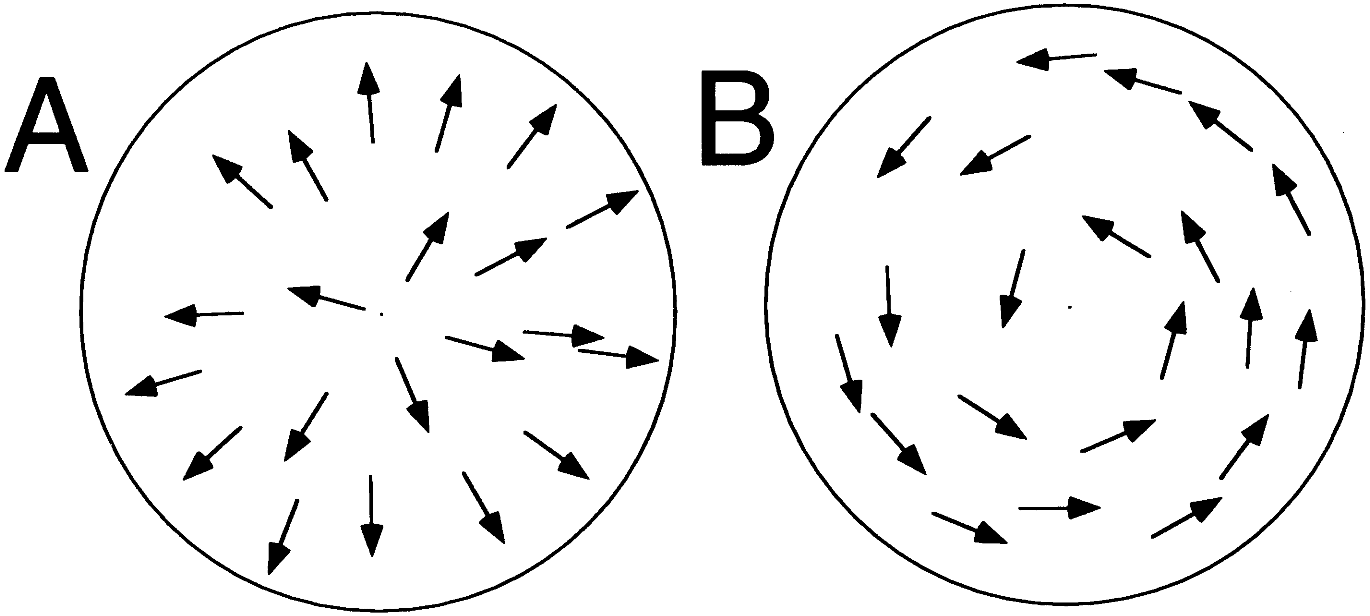


Figure 6

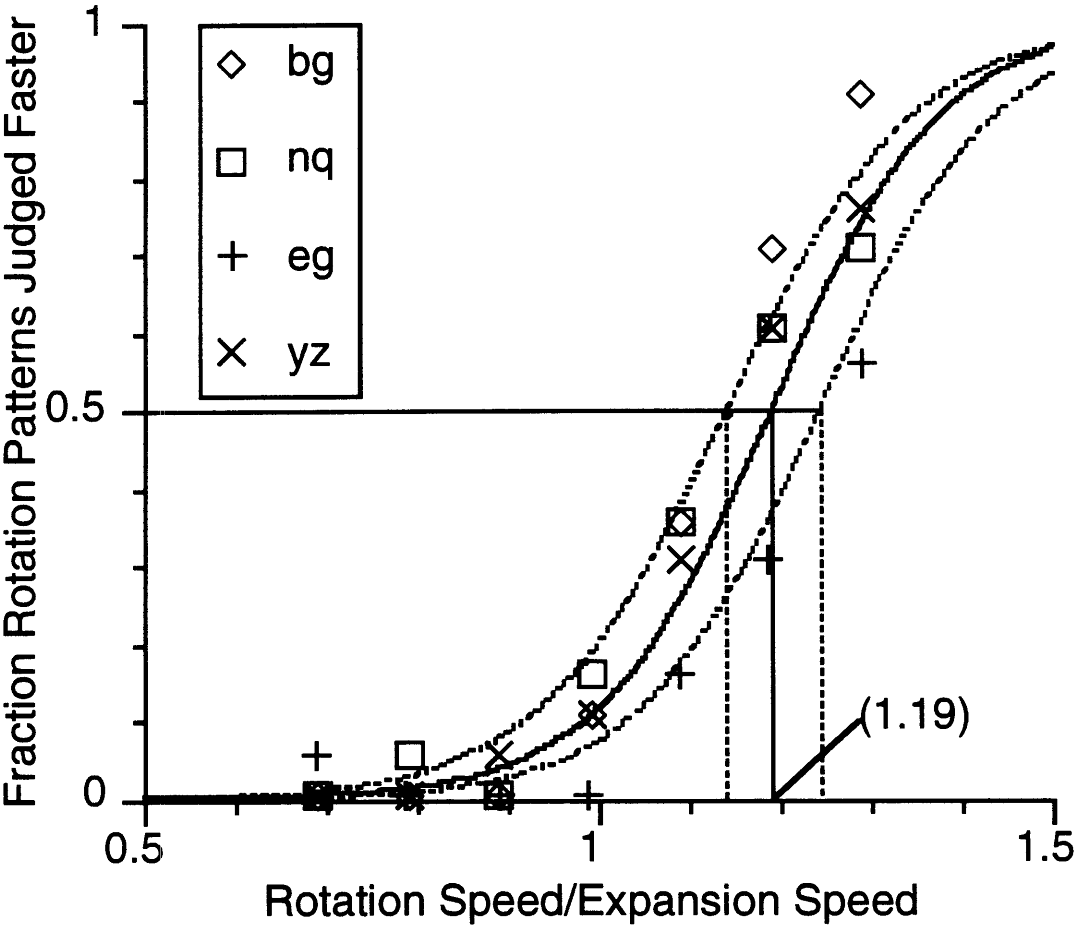

Fraction Rotation Patterns Judged Faster
1
0.5
0
0.5
1
1.5
Rotation Speed/Expansion Speed
◇ bg
□ nq
+ eg
× yz
(1.19)

Figure 7

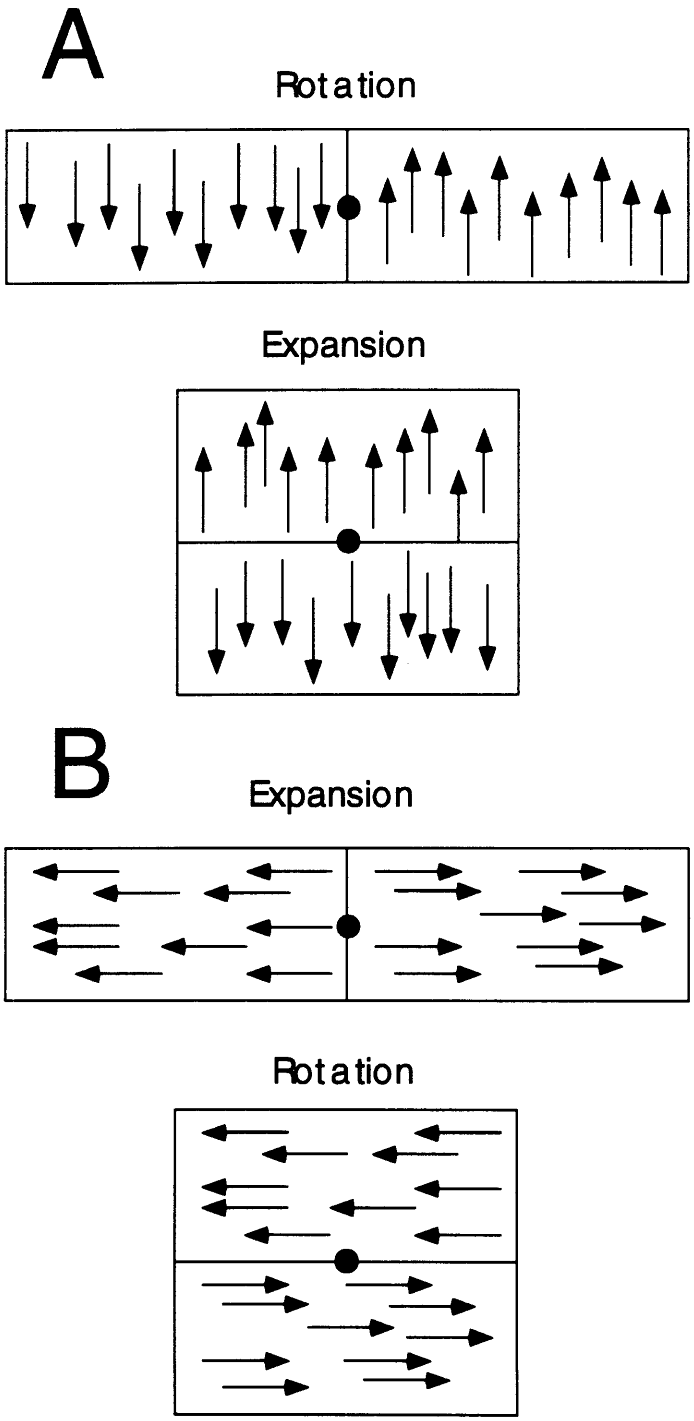


Figure 8

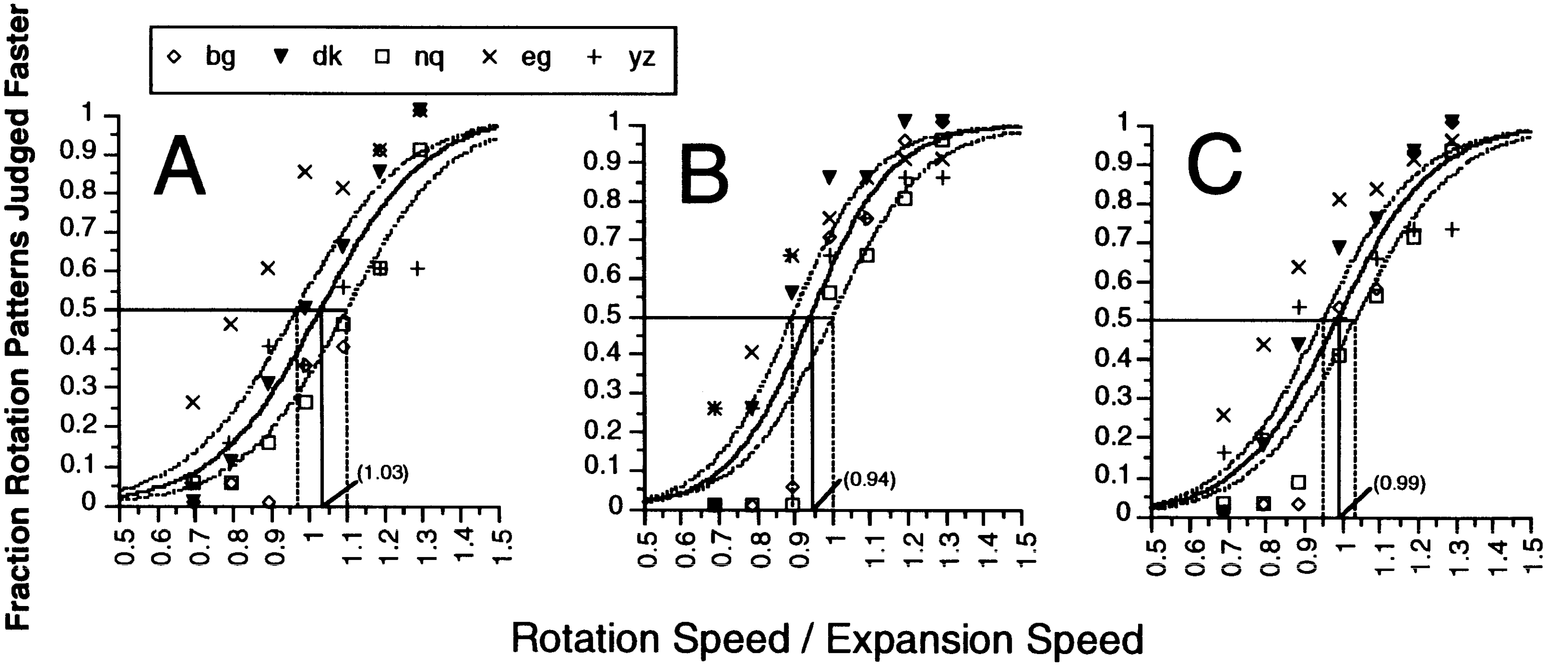
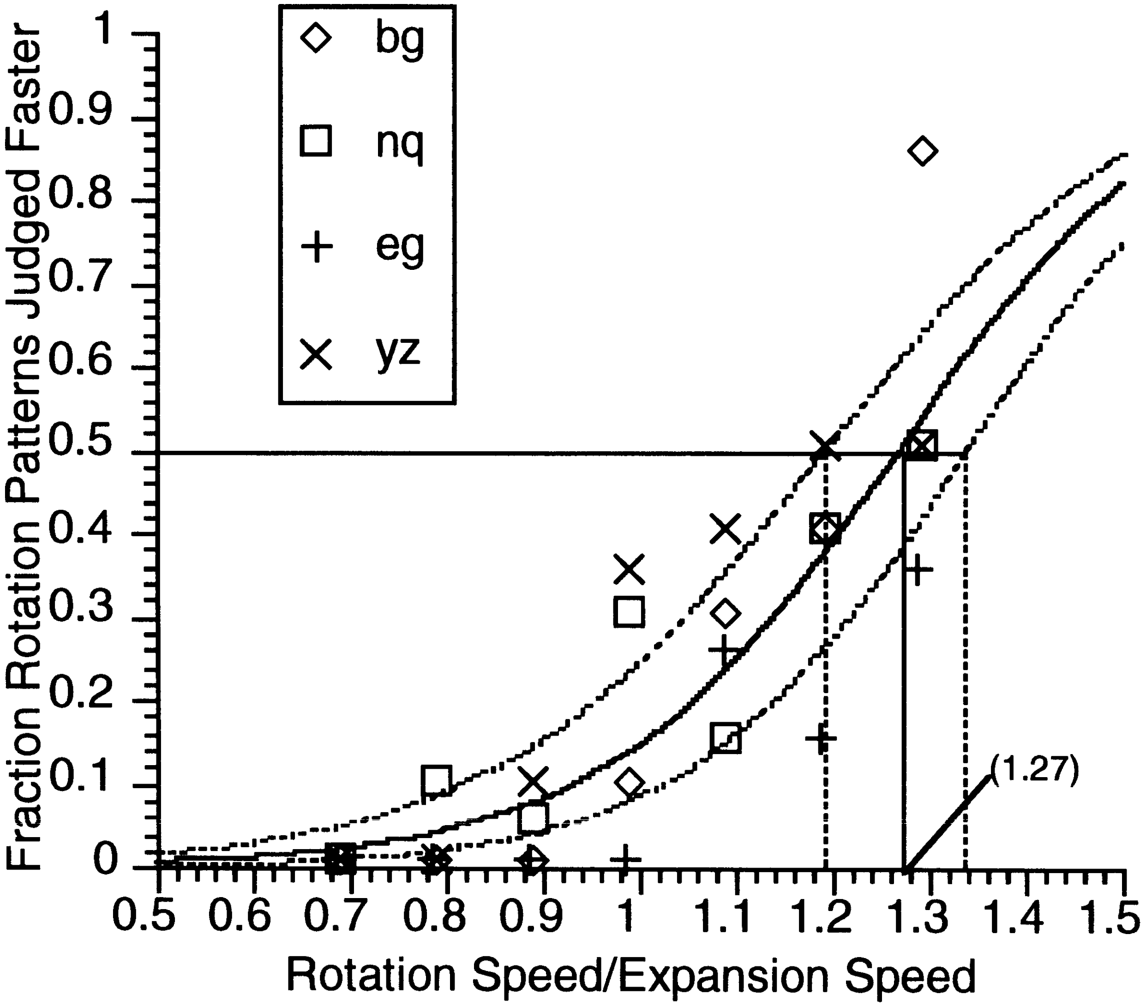

◇ bg ▼ dk □ nq × eg + yz
A
B
C
Fraction Rotation Patterns Judged Faster
Rotation Speed / Expansion Speed
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0.5
0.6
0.7
0.8
0.9
1
1.1
1.2
1.3
1.4
1.5
(1.03)
(0.94)
(0.99)

Figure 9



bg
nq
eg
yz
Fraction Rotation Patterns Judged Faster
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
(1.27)
0.5
0.6
0.7
0.8
0.9
1
1.1
1.2
1.3
1.4
1.5
Rotation Speed/Expansion Speed

Figure 10

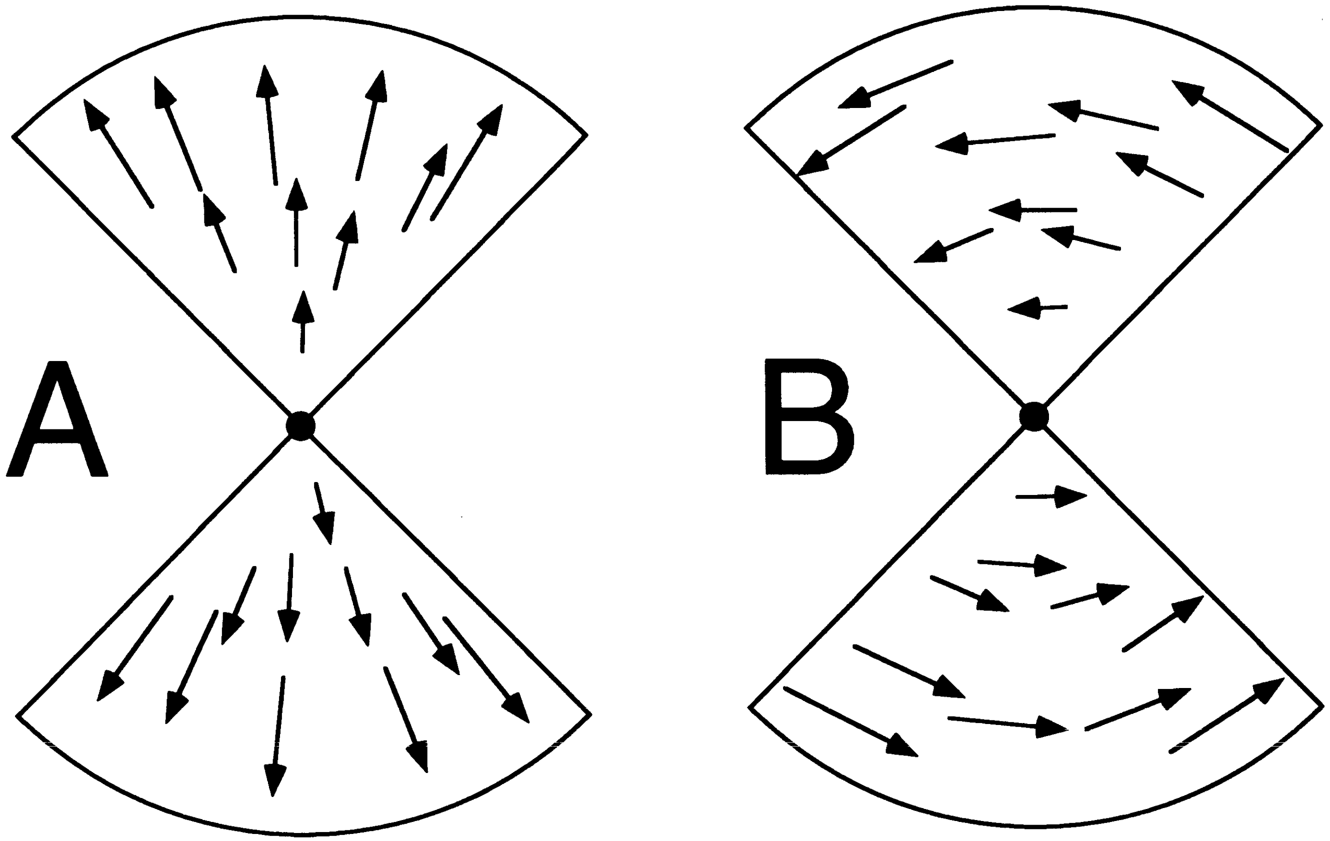


Figure 11

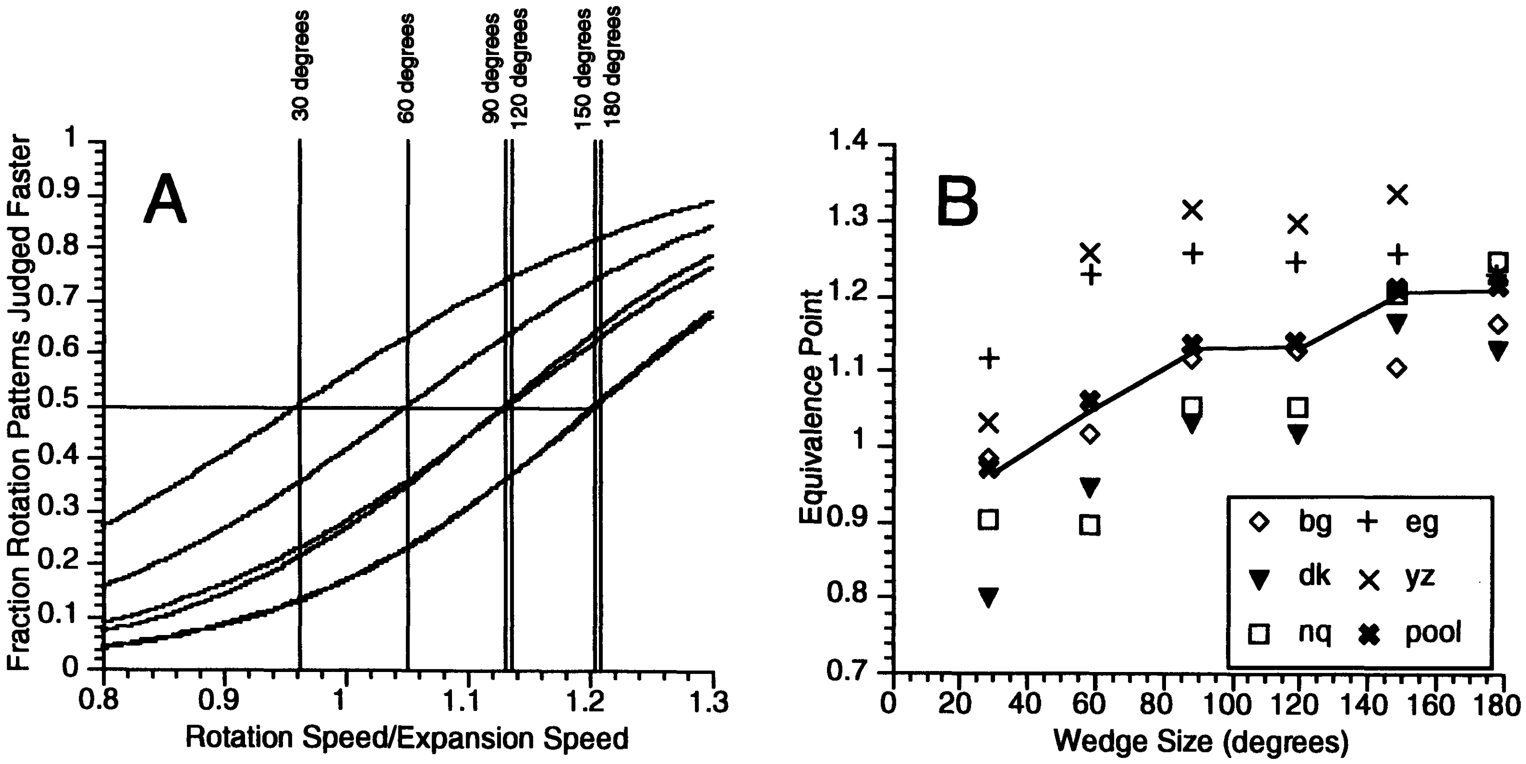

A
30 degrees
60 degrees
90 degrees
120 degrees
150 degrees
180 degrees
Fraction Rotation Patterns Judged Faster
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0.8
0.9
1
1.1
1.2
1.3
Rotation Speed/Expansion Speed
B
Equivalence Point
0.7
0.8
0.9
1
1.1
1.2
1.3
1.4
0
20
40
60
80
100
120
140
160
180
Wedge Size (degrees)
◇ bg
+ eg
▼ dk
× yz
□ nq
✖ pool

Figure 12

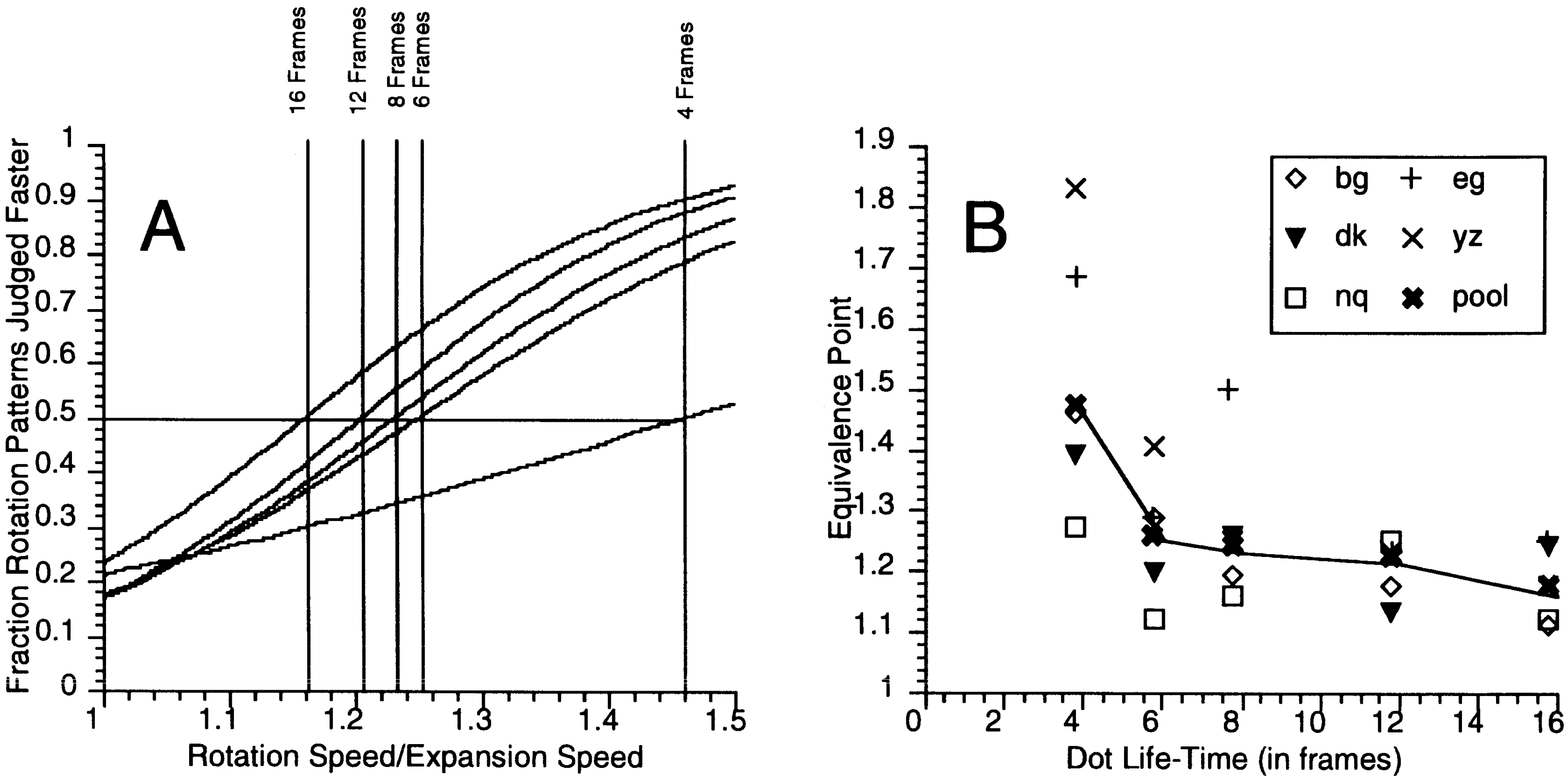
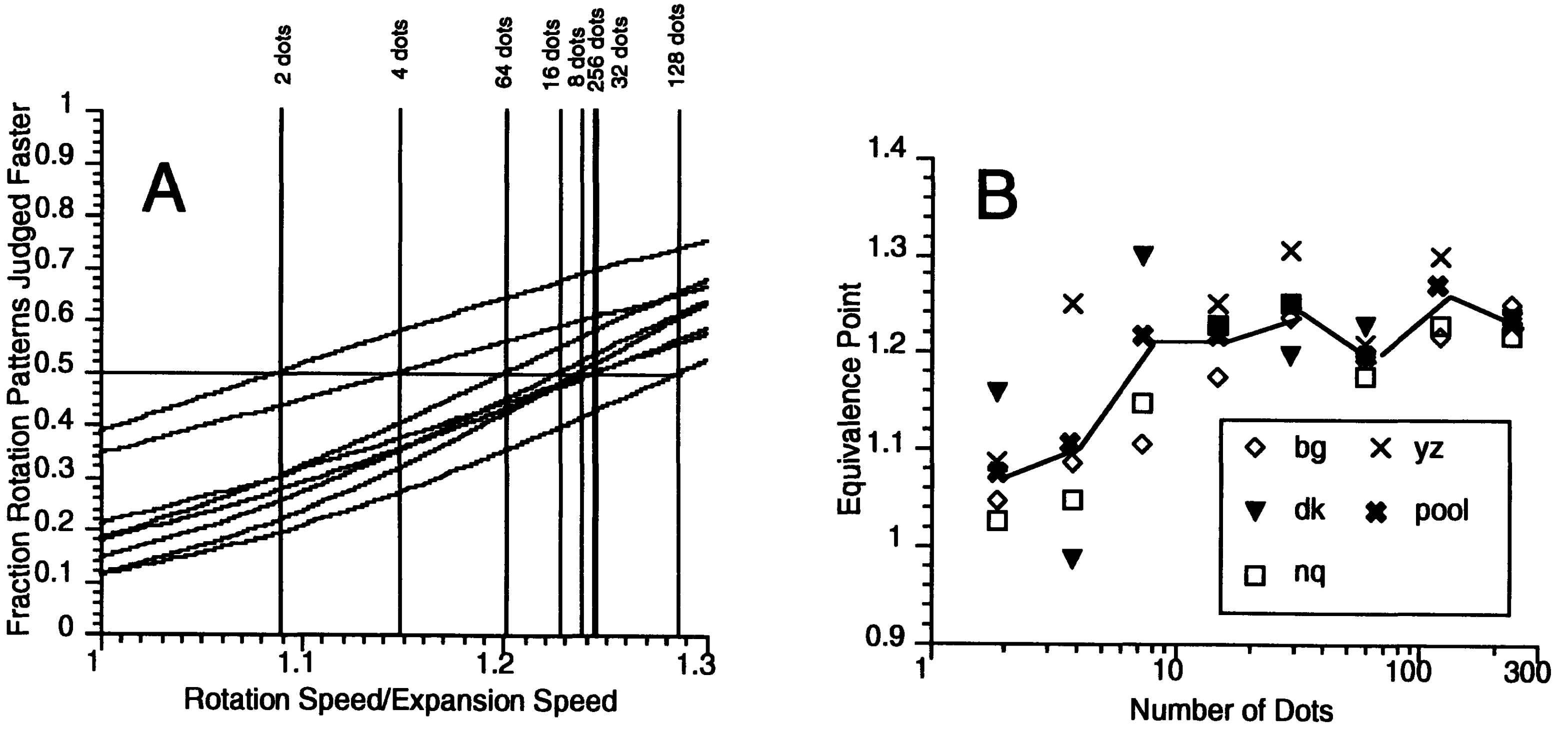

A
Fraction Rotation Patterns Judged Faster
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
1
1.1
1.2
1.3
1.4
1.5
Rotation Speed/Expansion Speed
16 Frames
12 Frames
8 Frames
6 Frames
4 Frames
B
Equivalence Point
1
1.1
1.2
1.3
1.4
1.5
1.6
1.7
1.8
1.9
0
2
4
6
8
10
12
14
16
Dot Life-Time (in frames)
◇ bg
+ eg
▼ dk
× yz
□ nq
✖ pool

Figure 13


A
Fraction Rotation Patterns Judged Faster
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
Rotation Speed/Expansion Speed
1
1.1
1.2
1.3
2 dots
4 dots
64 dots
16 dots
8 dots
256 dots
32 dots
128 dots
B
Equivalence Point
0.9
1
1.1
1.2
1.3
1.4
Number of Dots
1
10
100
300
◇ bg
× yz
▼ dk
✖ pool
□ nq