# Chapter 7

# Motion Processing in the Primate Cortex

RICHARD A. ANDERSEN
RALPH M. SIEGEL

**ABSTRACT**

*We developed a novel set of motion stimuli to test the ability of humans and monkeys to detect structure from motion. It was found, for all parameters tested, that humans and monkeys have similar abilities. Additional experiments tested the ability of monkeys and humans to detect the onset of shearing motion. Again, thresholds were found to be the same in the two species over a broad range of stimulus parameters. These two studies suggest that humans and monkeys have similar neural mechanisms for analyzing motion.*

*Small ibotenic acid lesions in monkeys trained in these motion detection tasks produced motion scotoma at locations in the visual field that corresponded to the locations of the lesions in the retinotopic representation of the middle temporal (MT) area. These deficits were specific for motion since contrast sensitivity thresholds were not affected. The shear-motion detection thresholds recovered within a few days. The structure-from-motion thresholds remained elevated after the motion detection thresholds recovered. These experiments demonstrate that area MT plays an important role in motion and structure-from-motion perception. The quick recovery of the motion detection thresholds raises the possibility that either area MT reorganizes itself after lesioning or parallel pathways that do not include area MT are involved in the recovery.*

This chapter discusses advances in our knowledge of how the brain analyzes moving stimuli. Results of anatomical, psychophysical, and recording/lesion experiments have elucidated important information regarding the way in which monkey and human brains process visual motion information. Of special importance, a presumed pathway for motion analysis has been anatomically pinpointed in the macaque monkey brain. This development is an important first step in understanding brain mechanisms for motion processing.

This motion pathway is hierarchical in structure, and an important first question is: What types of motion analysis occur at each level in this pathway? To address this question, we trained both human and primate subjects to carry out various psychophysical motion tasks. We then lesioned particular areas along the monkey's motion pathway and recorded the responses of neurons in these areas while the monkeys performed the tasks, and compared the results with the human subjects' performance on the same tasks. An equivalence of psychometric parameters between the two species would suggest that macaque and human brains use similar

motion-processing strategies and that the results derived from recording and lesion experiments in monkeys pertain to human visual motion processing as well.

First we review experiments evaluating the ability of monkeys to perceive motion and derive structure from motion. These were the first experiments to examine psychophysically the ability of monkeys to perceive visual movement (Golomb et al., 1985; Siegel and Andersen, 1987). Psychophysical studies of various motion detection abilities have, of course, been performed in humans for many years, but the stimuli used have included position as well as motion cues. We have therefore developed a class of novel motion stimuli that contain only motion cues, in order to isolate the motion-processing mechanisms of the brain. Since both the stimuli and the tasks were new, we also examined human performance on the tasks.

The second part of the chapter describes our physiological experiments, which we designed to examine the presumed functional hierarchy along the motion-processing pathway. Specifically, we discuss experiments exploring the effects of lesions at two stations in this pathway, the middle temporal (MT) area and the medial superior temporal (MST) area, on the monkeys' psychophysically determined abilities to perceive motion and derive structure from motion.

## ANATOMICAL PATHWAY

Experiments in the neocortex of monkeys have identified a presumed anatomical pathway for the analysis of visual motion. Elements of this pathway are shown in Figure 1. Direction-selective cells in the primary visual cortex (V1) project to MT, which is located in the posterior bank of the superior temporal sulcus and can be recognized by its characteristic heavy myelin staining. Recording experiments in this area indicate that most of its neurons are direction and speed selective, that is, tuned to respond differentially depending on the direction and speed of movement (Van Essen, 1985; Maunsell and Newsome, 1987). These observations have led to the suggestion that MT is specialized for motion analysis (Zeki, 1974). However, the role of MT in motion perception has until recently not been directly tested by measuring the effects of lesions in MT on the ability of animals to perceive motion (Newsome and Pare, 1986; Siegel and Andersen, 1988).

MT in turn projects to areas in the inferior parietal lobule, including MST, which is located in the anterior bank of the superior temporal sulcus, and the lateral intraparietal (LIP) area, which is located in the lateral bank of the intraparietal sulcus. Recordings from MST indicate that its cells typically have large bilateral receptive fields and many of its cells have responses that are selective for expanding or rotating velocity fields (Saito et al., 1986; Sakata et al., 1986; Tanaka et al., 1986). These intriguing experiments suggest that MST may be specialized for the processing of more complex aspects of motion such as structure from motion, whereas MT processes simpler aspects of motion such as direction and speed, and also that there is a functional hierarchy in motion processing along the V1-to-MT-to-MST pathway. Unfortunately, position as well as motion cues were present in the rotation and expansion stimuli used by previous investigators. Recording experiments using


**Figure 1.** *Coronal section through a macaque monkey brain containing a large part of the visual motion-processing pathway in the extrastriate cortex.* This section has been stained for myelin. The arrows represent borders of the different cortical areas. MT and LIP can be recognized by their characteristic deep myelin staining. MT, medial temporal area; MST, medial superior temporal area; LIP, lateral intraparietal area. (From Andersen et al., 1990.)

more highly controlled stimuli will be necessary in order to establish precisely the role of MST in the processing of structure from motion.

LIP, one of the other areas receiving ascending projections from MT, has recently been shown to play a role in the processing of saccadic eye movements (Gnadt and Andersen, 1988). Why this area receives inputs from the motion-processing pathway is unclear. One possible reason for this saccade area to receive inputs from the motion pathway is that the motion information may be used to calculate the correct amplitude for saccades made to moving targets.

The elucidation of the anatomical location of this motion-processing pathway has been an important first step in understanding how the brain processes the perception of motion. Part of our experimental approach has been to develop novel motion stimuli that do not contain position cues and to use them to analyze the ability of monkeys and humans to detect motion and structure from motion. We

have reasoned that similar thresholds in psychophysical motion tasks obtained from these two species would suggest that their cortical motion-processing machinery is essentially the same and that information gained from physiological experiments in monkeys, described below, would be applicable to humans.

## PSYCHOPHYSICS

Motion stimuli can be divided into two large classes, those that produce absolute motion and those that produce relative motion. Absolute motion is movement of a particular velocity that is constant for all points in the visual field, whereas relative motion results from movements of different velocities at different locations in the visual field. Absolute motion rarely occurs in natural situations, the one exception being the movements of the retinal image generated by eye movements. Relative motion results from movement of an object against a background or locomotion through the environment by the observer. Relative motion provides useful information to the organism such as depth perception based on motion parallax, the separation of figure and ground, and time to collision based on looming cues.

The human psychophysics of relative motion perception have been studied for many years using discrete moving objects. A major problem with this class of stimuli is that it contains position as well as motion cues. A good example of this position contamination problem was given by Nakayama and Tyler (1981) when they described the perception of motion of the minute hand of a clock. Because it moves slowly, it is difficult to perceive its motion. In a motion detection task one cannot determine if it is the actual motion of the hand or the change of its position on the clock face that is perceived.

Investigators have begun using moving random dot patterns as relative motion stimuli. It is argued that these stimuli are complex and it is therefore more difficult to memorize the dot patterns to determine whether they change shape on the introduction of the motion stimulus. However, these stimuli still have position cues and subjects have shape change cues as well as motion cues available to them. In the experiments discussed below we used moving random dot displays to which we introduced finite lifetimes. As a result, these "dynamic" patterns are constantly and randomly changing shape, thus finally eliminating shape change as a cue.

When we first began these experiments, no one had measured the ability of monkeys to see motion by psychophysical experiment. We chose to examine two types of relative motion perception in monkeys, the ability to detect shear motion and the ability to detect two- and three-dimensional structure in velocity fields. We also examined the spatio-temporal integration characteristics of structure-from-motion perception. These experiments suggest that the brain forms neural representations of surfaces using structure-from-motion information.

## SHEAR THRESHOLDS

The first type of motion perception we studied in monkeys performing psychophysical tasks was the detection of shearing motion (Golomb et al., 1985). Shearing motion is a class of relative motion in which the change in direction of motion

occurs along the axis orthogonal to the direction of motion. This form of motion is of great physiological importance as it accounts for depth perception based on motion parallax and is used to differentiate foreground and background using motion discontinuities. Recording experiments in monkey MT (Allman et al., 1985), the pigeon optic tectum (Frost and Nakayama, 1983), and monkey area V2 (Orban et al., 1986) have identified neurons with receptive fields that have opponent direction center–surround organization; these cells would be maximally active for a shear stimulus. Because of its importance, a great deal of psychophysical research into shear motion detection abilities has already been done on human subjects (Rogers and Graham, 1979, 1982; Nakayama, 1981; Nakayama and Tyler, 1981; Nakayama et al., 1985), making it possible to compare our results on monkeys and humans to those of previous work.

The stimulus we used is shown in Figure 2. The animals and humans performed a reaction time task in which they had to release a key to indicate the detection of the onset of shear motion in a static random dot pattern. Both static and dynamic (finite point life) displays have been used with similar results. The shear motion consisted of a standing transverse wave of sinusoidally varying spatial and temporal frequencies. Thresholds for shear detection were measured for various combinations of spatial frequency, temporal frequency, and amplitude. Figure 3 shows data from two monkeys and two humans in which amplitude thresholds were determined at different spatial frequencies. The temporal frequency was held constant at 2 Hz, the optimum temporal frequency for both species. Several interesting results are illustrated in this figure: (1) Both the humans and the monkeys showed characteristic U-shaped curves with the minimum threshold at spatial frequencies between 0.1 and 0.6 cycles per degree. The amplitude thresholds for humans and monkeys were consistently found to be 13–15 arc-sec for 20% thresholds (performance just above chance in the task) and 22–25 arc-sec for 50% thresholds (correct responses in 50% of the trials). Human subject BG had markedly lower thresholds than the two monkeys and four other humans tested, with a 20% threshold at 4 arc-sec and 50% threshold at 7 arc-sec, indicating that there is a range of thresholds among subjects. (2) The amplitude thresholds are extremely small, particularly when one considers that they are peak-to-peak amplitudes separated by 5° at a spatial frequency of 0.1 cycles per degree. These thresholds are similar in range to the hyperacuity thresholds measured in Venier acuity tasks. (3) The spatial frequency tuning curve was found to be very "low pass," that is, it allowed the detection of only low spatial frequencies. In fact, this curve is an octave lower than the curve for contrast sensitivity. These different spatial frequency threshold curves suggest that there are different neural pathways for motion and contrast perception.

## STRUCTURE-FROM-MOTION PSYCHOPHYSICS

We developed a novel set of stimuli that test the ability of subjects to detect structure in velocity fields without introducing confounding positional cues. We examined three types of structures in both monkeys and humans (Siegel and Andersen, 1987): The two two-dimensional (2-D) structures are expansion and rotation (Figure 4), while the three-dimensional (3-D) structure is a revolving hollow cylinder

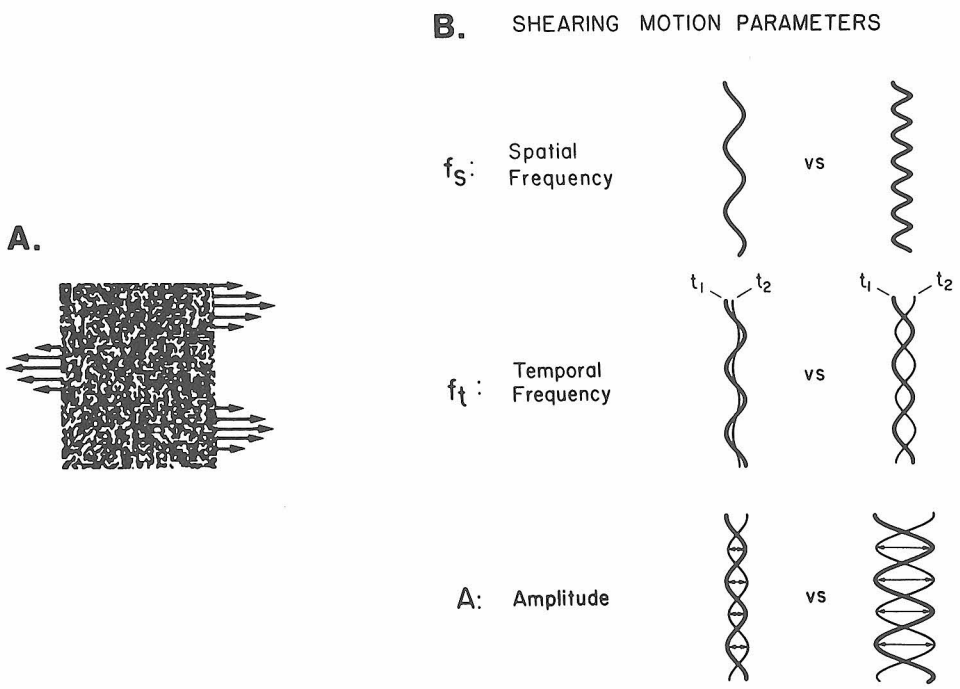



**Figure 2.** *Shearing motion stimulus. A:* Random dot display that undergoes horizontal shearing motion. Each horizontal row of dots moves as a rigid unit with a velocity that is a sinusoidal function of the vertical position of the row. The arrows represent the instantaneous velocity vectors. Note that the direction of motion reverses at the zero crossings of the sinusoidal velocity function. *B:* Displacement profiles that schematize the parameters of motion used in this study. The spatial frequency parameter ($f_s$ = cycles per degree) determines the vertical distance between the (nonnodal) random dot rows moving with the same instantaneous velocity vector. The function of sinusoidally varying velocities diagramed on the left is of a lower spatial frequency than the one on the right. The temporal frequency ($f_t$ = hertz) parameter determines the speed of peak-to-peak movement of the random dot rows. The displacement profiles illustrating this point have the same spatial frequency and amplitude as modulation parameters; however, the temporal frequency of modulation is less for the example on the left. As a result, the dots in the display with the parameters on the left traverse a smaller distance between times $t_1$ and $t_2$ than they do in the display with the parameters on the right. Amplitude represents the peak-to-peak distance traversed by the horizontal rows of random dots corresponding to the antinodes of the driving spatial frequency. The functions in the figure representing the amplitude parameter show two amplitudes of modulation for motion displays having the same spatial frequency and temporal modulation characteristics. (From Golomb et al., 1985.)

(Figure 5). The standard number of points tested was 128. The subjects performed a reaction-time task in which they had to detect the transition from an unstructured velocity field to a partial or completely structured velocity field. The structured and unstructured displays contained the same vectors, the only difference being that in the unstructured displays the vectors were shuffled randomly to destroy structure.

Figure 4, the rotation case, explains this structure-from-motion task better. The subject first viewed the noise field in the lower right-hand corner. At some random time between 1 and 4 sec, the display changed to a partially (Figure 4, middle right-hand panel) or completely (Figure 4, upper right-hand panel) structured rotation. The subject had to release a key quickly when he detected this transition. If the

subject's response fell within a reaction time window just after the transition, he received positive feedback in the form of a sound (humans) or a drop of juice (monkeys). Since it is more difficult to see a transition to a lesser degree of structure, we derived the psychometric curves by varying the degree of structure in the test stimulus and measuring the number of correct responses.

Figure 6 demonstrates, for the rotation case, how the displays were destructured. The computer first generated the correct 2-D pattern (Figure 6, left-hand panel). However, before the velocity field was displayed each vector was displaced on the surface of the screen by a random amount not to exceed a certain upper limit. If this limit was the entire width of the screen, the display was assigned a structure value of 0.0 (Figure 6, right-hand panel). If, on the other hand, the upper limit was half the screen's width, then the display had a 0.5 fraction of structure (Figure 4, middle right-hand panel) and rotation could be perceived in such a display. The method described above for rotation was also used for the expansion case and the 3-D cylinder (Figure 7).

In all displays the moving dots had finite point lives, typically 270 msec in duration. Although a constant number of points was always present on the screen and each had the same lifetime, the points were born at random times with respect to one another and appeared at random locations on the screen. Finite point lives are important for three reasons: (1) They introduce continuous random change to local shapes in the display. Thus any local shape change introduced by a change in the structure of the velocity field is different from trial to trial and is indistinguishable from the other, constantly occurring, shape changes. (2) Finite lifetimes allow

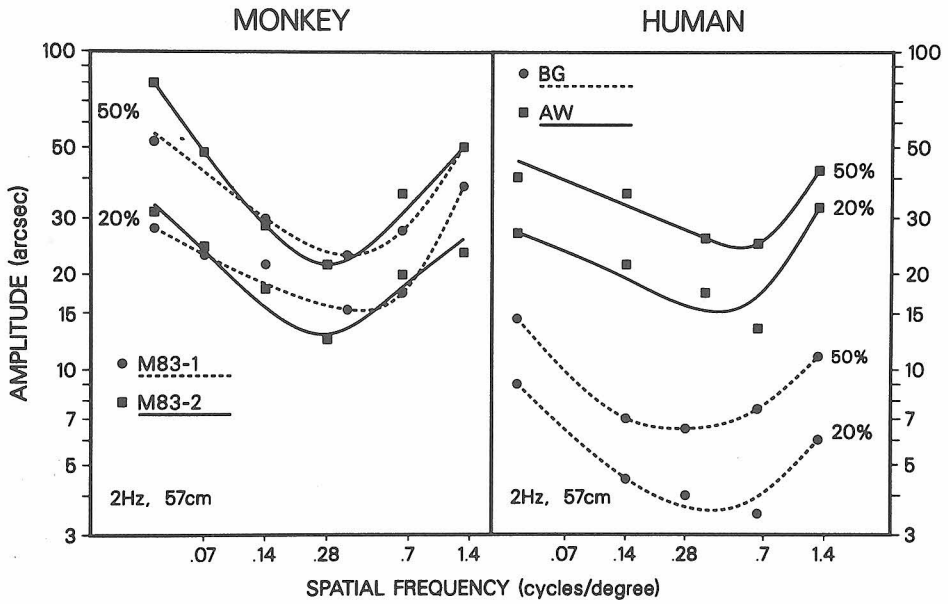



**Figure 3.** *Comparisons of spatial frequency profiles at 57-cm viewing distance for monkeys (M83-1 and M83-2) and humans (AW and BG); 50% hit values are shown.* Both monkey profiles fell within the range defined by the human curves. All the subjects did best at spatial frequencies of around 0.14–0.65 cycles per degree, demonstrating both low- and high-frequency upturns in threshold. (From Golomb et al., 1985.)

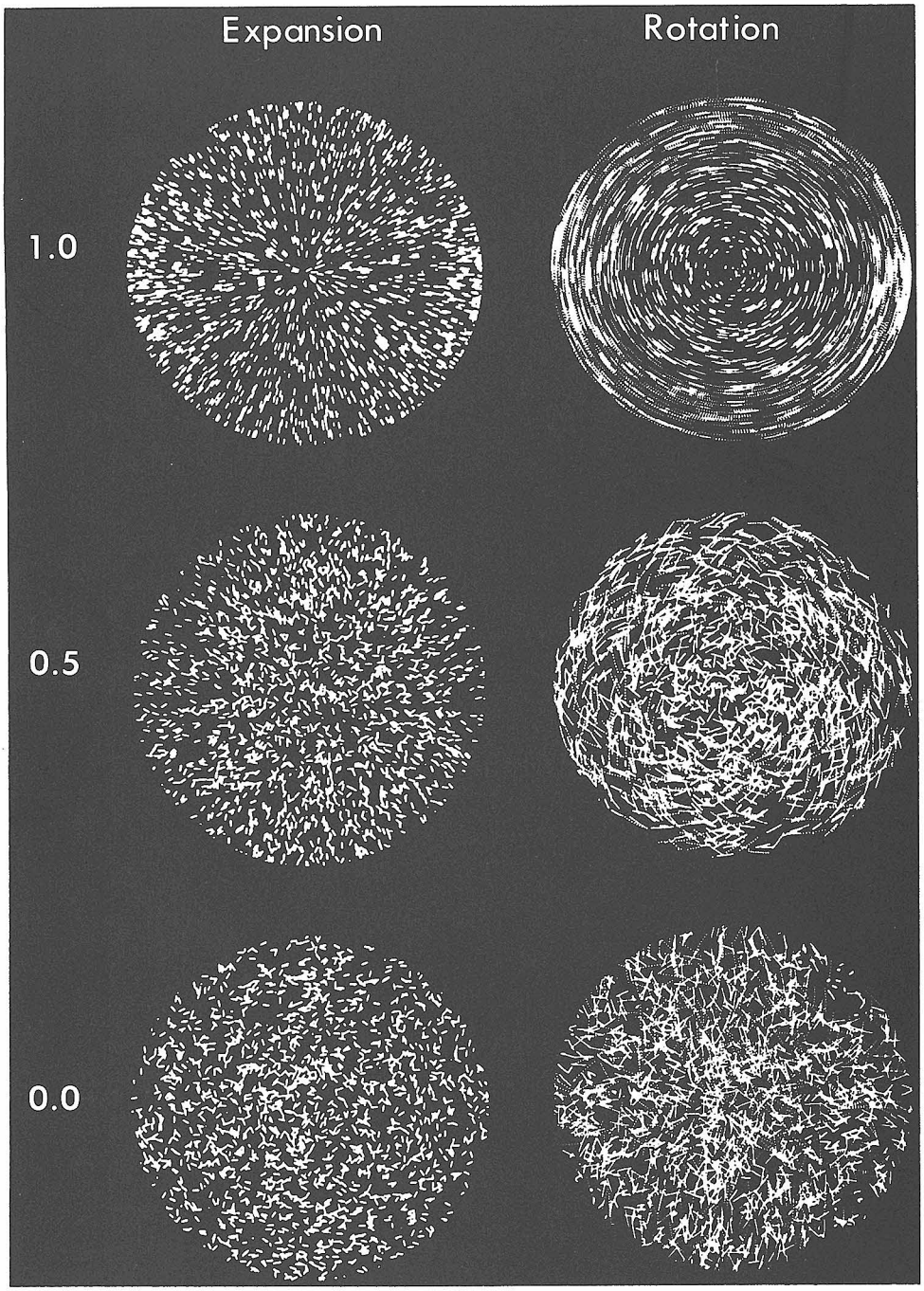



**Figure 4.** *Time-lapse photograph of the 2-D expansion and rotation stimuli.* The top displays (1.0) are completely structured. The middle displays have a fraction of 0.5 structure and the bottom displays (0.0) are noise.

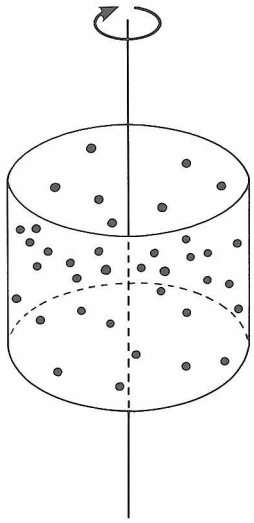


**Figure 5.** *3-D structure-from-motion stimulus.* Dots are projected onto the surface of a revolving hollow cylinder.

for smooth transitions to structure. The structure is changed only for new points, and as a result complete transition requires an amount of time equal to the lifetime of the points. (3) In the 3-D cylinder task, if the dots are on continuously they become denser at the edges of the display due to the orthographic projection and result in a density cue in the structure transitions. However, with finite lives the dots are randomly and evenly plotted and replotted across the screen and do not accumulate at the edges.

Figure 8 shows psychometric functions for the 3-D cylinder detection task for a monkey and a human, and it can be seen that the psychometric functions for the two species are very similar. Such equivalent performance was also seen in monkeys and humans for both of the 2-D detection tasks.

It was important to determine whether the subjects were using global cues from the entire flow field or simply local cues such as local changes in the speed or direction of the dots to detect the emergence of the test structures. To make this

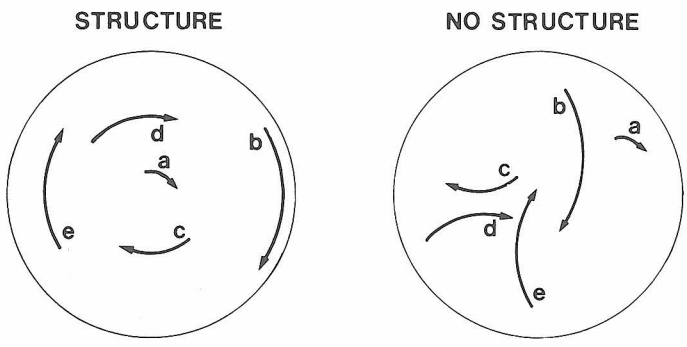



**Figure 6.** *Structured and unstructured displays for the 2-D rotation stimulus.*

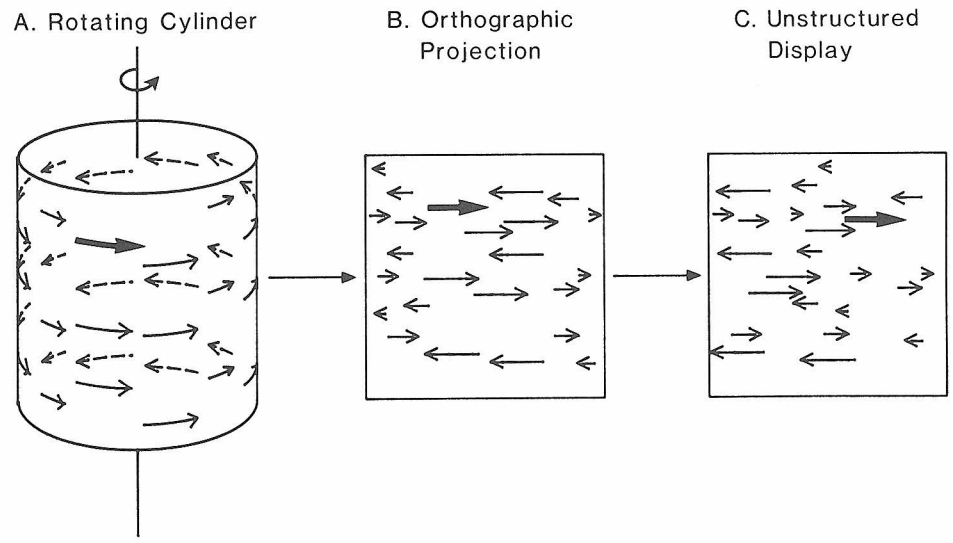



**Figure 7.** *Method of generating the structured and unstructured 3-D revolving hollow cylinders.* (From Siegel and Andersen, 1987.)

distinction, the field was masked so that subjects could see only a small fraction ($\frac{1}{9}$ or $\frac{1}{25}$) of the display and thus had to perform the task using only local cues. Figure 9 shows the psychometric curves for a human subject performing the task with and without the masks; the subject could not perform above chance levels when the mask was present, even when the transition was to 100% structure. The same results were found for the monkey with the 3-D cylinder, and both monkeys and humans also showed raised thresholds for the two 2-D tasks. Thus it can be con-

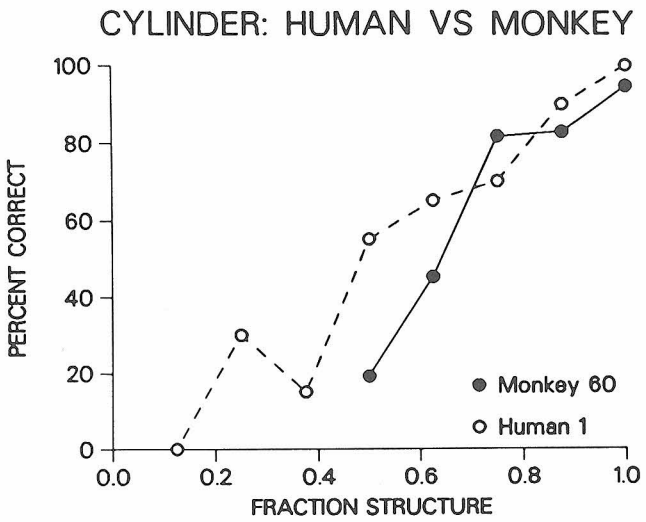



**Figure 8.** *Psychometric function for a monkey and a human using the 3-D revolving cylinder task.* The fraction of structure is plotted on the abscissa and the percentage of correct responses on the ordinate. Note the similarity of the two psychometric functions.

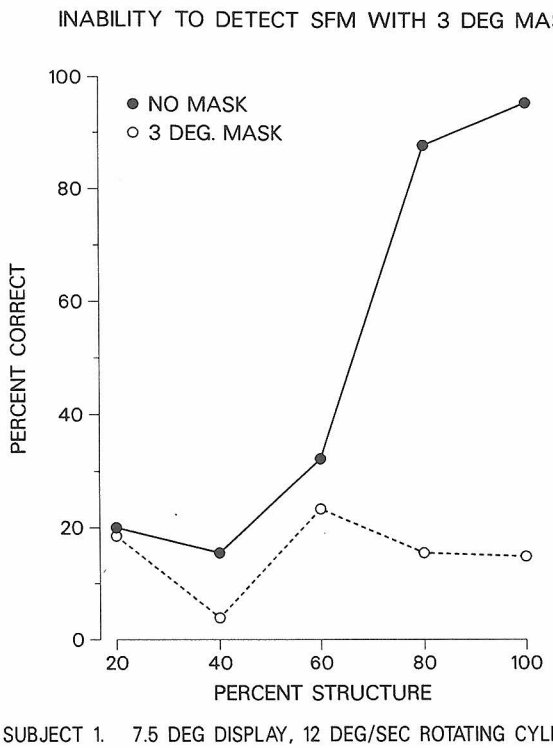



**Figure 9.** *Local versus global processing in the 3-D cylinder task.* The continuous line graph shows the psychometric function of a human viewing the entire surface of a 7.5° cylinder. The dotted line shows the performance of the subject when all but a 3° × 3° square in the center of the display is masked. Note that the subject cannot do the task, even at the transition to complete structure (100%), when the display is partially masked. Similar results were found for the monkey subjects.

cluded that the subjects use information integrated over the entire flow field to perform these tasks.

A second, more complex question is whether the subjects use neural representations of the 3-D shape of the cylinder to perform the task or only the cues that are present in the 2-D velocity field. For instance, the 2-D velocity field for the cylinder has sinusoidally varying speeds across its surface, with slow speeds at the edges and fast speeds at the center. This 2-D distribution of speeds might serve as a perceptual cue. Unfortunately, the 2-D distribution cannot be altered for control purposes as it would destroy the 3-D percept. This problem has plagued human psychophysical experiments for some time and is even more acute in monkey experiments, where the animal cannot verbally report its perception of 3-D structure.

From the reaction time data, however, we now have indirect evidence that both monkeys and humans in fact use 3-D percepts of the cylinder to perform the structure-from-motion task. On average it requires monkeys and humans 200 to 400 msec longer to do the 3-D cylinder task than the 2-D rotation task. The 2-D cues in the 2-D and 3-D displays are similar but inverted. The 2-D rotational flow field is slow in the center and fast at the edges, whereas the 3-D cylinder flow field is fast in the center and slow at the edges. If the subject were using the 2-D flow field to perform the cylinder task, one would expect the reaction time for the rotation and

cylinder stimuli to be nearly the same. The fact that the 3-D task requires a dramatically longer reaction time suggests that the brain has to compute the 3-D structure of the flow field and that this computation requires more time than the 2-D computation. An alternate view is that the longer reaction time is due to the 3-D task being more difficult in some general way. This argument may not, however, be tenable as preliminary experiments have demonstrated that changing features of the display, such as the lifetime of the dots, change reaction times. However, the reaction time difference between the 2-D and 3-D tasks remains approximately constant (M. Husain and R. A. Andersen, unpublished observations). Another alternative we considered is the possibility that because the 3-D cylinder is hollow and has two surfaces sliding past each other in opposite directions, perhaps the increased reaction time is due to the time required for the perceptual segregation of the two surfaces. To test this possibility, we modified the 2-D rotation task to include two superimposed and opposite 2-D rotations so that the 2-D task would also require the segregation of surfaces moving in opposite directions. This two-surface rotation did not increase reaction time over that of the one-surface rotation. Thus the reaction time data provide indirect evidence that the subjects performing the 3-D detection task use 3-D perception of the surface of the cylinder.

## SPATIAL AND TEMPORAL INTEGRATION

Having established in our control experiments that subjects were detecting the structure of the velocity fields, we were ready to examine how information about structure from motion is integrated by the visual system over space and time. Ullman (1979) laid the important theoretical groundwork for this problem with his structure-from-motion theorem that "given three distinct orthographic views of four non-coplanar points in a rigid configuration, the structure and motion compatible with the three views are uniquely determined." He recognized that this was simply an algebraic limit and that the brain probably requires more points and frames since it is a so-called "noisy" system and may also use a less than optimum algorithm. We were then in a position to address this issue directly using our controlled stimuli. We investigated the question of spatial integration by varying the density of the points and the temporal integration problem by varying the point lifetimes in the 3-D cylinder task.

Figure 10 shows the performance of a monkey and a human subject for different point lives with the standard 128 points present in the display; below 100 msec the human subject performed poorly and below 50 msec the monkey did poorly. For a fixed point life of 532 msec the human and monkey subjects both showed poor performance when there were fewer than 32 points. Interestingly, there was a space–time tradeoff so that equal performance was achieved with longer point lifetimes and fewer points or vice versa. Thus these experiments show that movement information is integrated in both space and time to form neural representations of 3-D surfaces.

### Surfaces Are Recovered From Motion by the Visual System

Our space–time integration experiments offer insights into the neural mechanisms for computing structure from motion. One possible algorithm describes the brain as

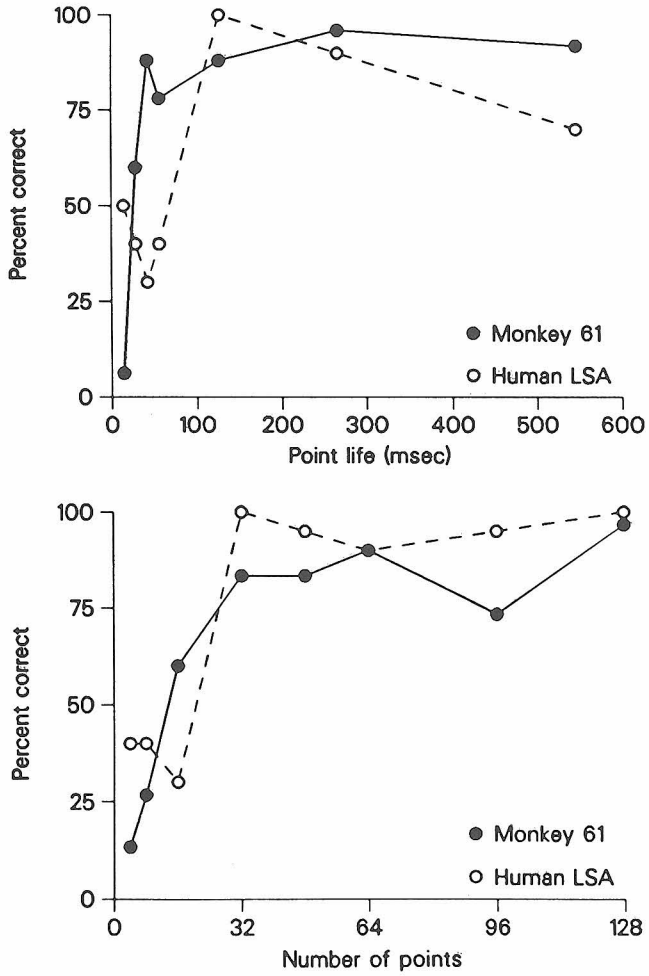



**Figure 10.** *Examples of time (A) and space (B) integration for the 3-D cylinder task.* The upper plot shows performance as a function of point life. The display had 128 points and the test stimulus had a 0.875 fraction of structure. Point lives of less than 100 msec reduced performance for both the monkey and human subject. The lower panel shows performance as a function of the number of dots in the display. Point life was held constant at 532 msec and the number of points was varied. Both species performed poorly when there were fewer than 32 points. In both cases the refresh rate was 70 Hz and the angular velocity 35°/sec. (From Siegel and Andersen, 1987.)

tracking the location of individual points in absolute coordinates continuously to determine whether their trajectories conform to the movement of a rigid object (Ullman, 1984). In this algorithm the brain solves algebraic equations using the points in the image. However, we know from our experiments that this cannot be the case. The reaction time data show that the 3-D cylinder structure-from-motion computation required several hundred milliseconds, and thus each dot would have to be tracked over a long period of time. However, subjects performed well on the time integration experiments even when the dots were visible for only 100 msec and were then replotted at new, random positions, thus making any continuous, point-bound computation impossible. These results suggest that a surface is computed using a large number of intermittent sampling points across that surface.

Other aspects of the experiment suggest that the brain extracts surfaces. The demonstration of the masking experiments that the subject uses information from the entire velocity field suggests that the entire surface is computed. Also, the different reaction times on the 2-D and 3-D tasks indicate that computation time is a function of the type of surface being computed. Finally, the percept obtained by human observers from these displays is of a shape with a definite surface.

These results suggest that there is one subsystem in the motion-processing pathway that is important in extracting surfaces. There may also be a second motion subsystem that extracts edges, as indicated by motion discontinuities. These discontinuities are generated by shear motion or occlusion, and cells have been found in areas V2, MT, and the superior colliculus that respond to these motion-generated borders (Frost and Nakayama, 1983; Allman et al., 1985; Orban et al., 1986). Allman (1987) has proposed that two systems exist in the contrast system of area V1, one that extracts borders and the other that extracts surfaces. The cytochrome oxidase–rich "blob system" comprises the surface system and includes cells that are generally not orientation tuned, are selective for low spatial frequencies, and are color selective, color generally being a quality of surfaces. The "interblob" area, on the other hand, is made up of cells that are orientation tuned, respond to high spatial frequencies, and are not as color selective as those of the blob region. An argument against their functioning exclusively in color perception comes from the observation that even nocturnal New World monkeys, which have no color vision, have well-developed blob systems. It would be interesting if a similar functional segregation were found in the motion system, with separate analyzers for motion generated by surfaces and by edges. Would such a system coexist with the presumed surface and edge areas of the contrast system?

Ullman (1984) has proposed an algorithm for computing structure from motion that assigns structure to points in the image. Then a new frame is selected and the structure is changed by the smallest amount compatable with the motion of a rigid object between frames. This process is reiterated until a stable shape is achieved. Hildreth and colleagues (Grzywacz et al., 1986; Inada et al., 1986) have shown that this incremental rigidity algorithm accounts for many observations from their human psychophysical studies. This algorithm requires the use of the absolute positions of points tracked over all frames, which is not compatable with our results for limited point lives. However, with modification this algorithm can most likely be used to compute surfaces using intermittent sampling from flow fields (E. C. Hildreth, personal communication).

## MIDDLE TEMPORAL AREA LESIONS

Having psychophysically determined many aspects of the motion-processing abilities of monkeys, we are now in the position to begin examining the neural machinery that accounts for motion perception, using these psychophysical techniques in conjunction with lesion and recording techniques. Our first set of experiments examined the effects of restricted lesions in MT on the ability of monkeys to see shear motion and structure from motion (Siegel and Andersen, 1986). As mentioned in the introduction, several lines of recording evidence have indicated that MT plays a role in motion perception. However, this hypothesis had not been directly tested by assessing the effects of MT lesions on various motion detection abilities.

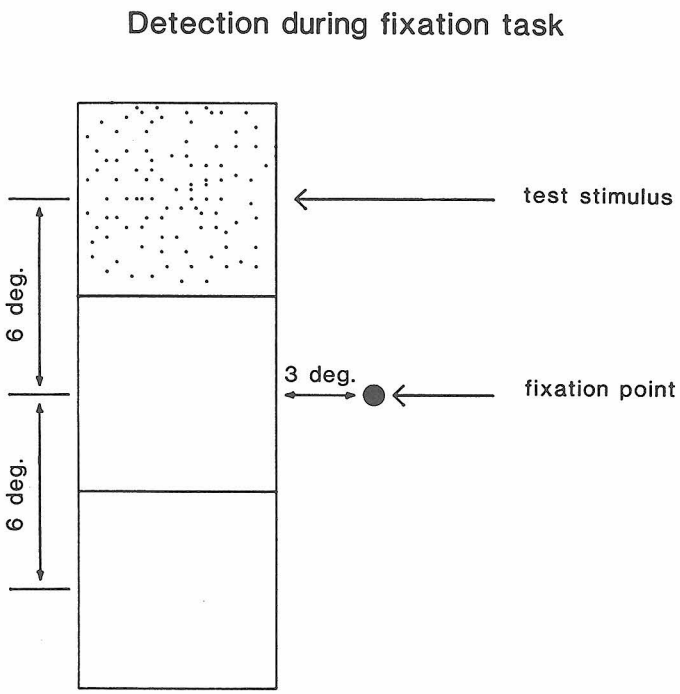


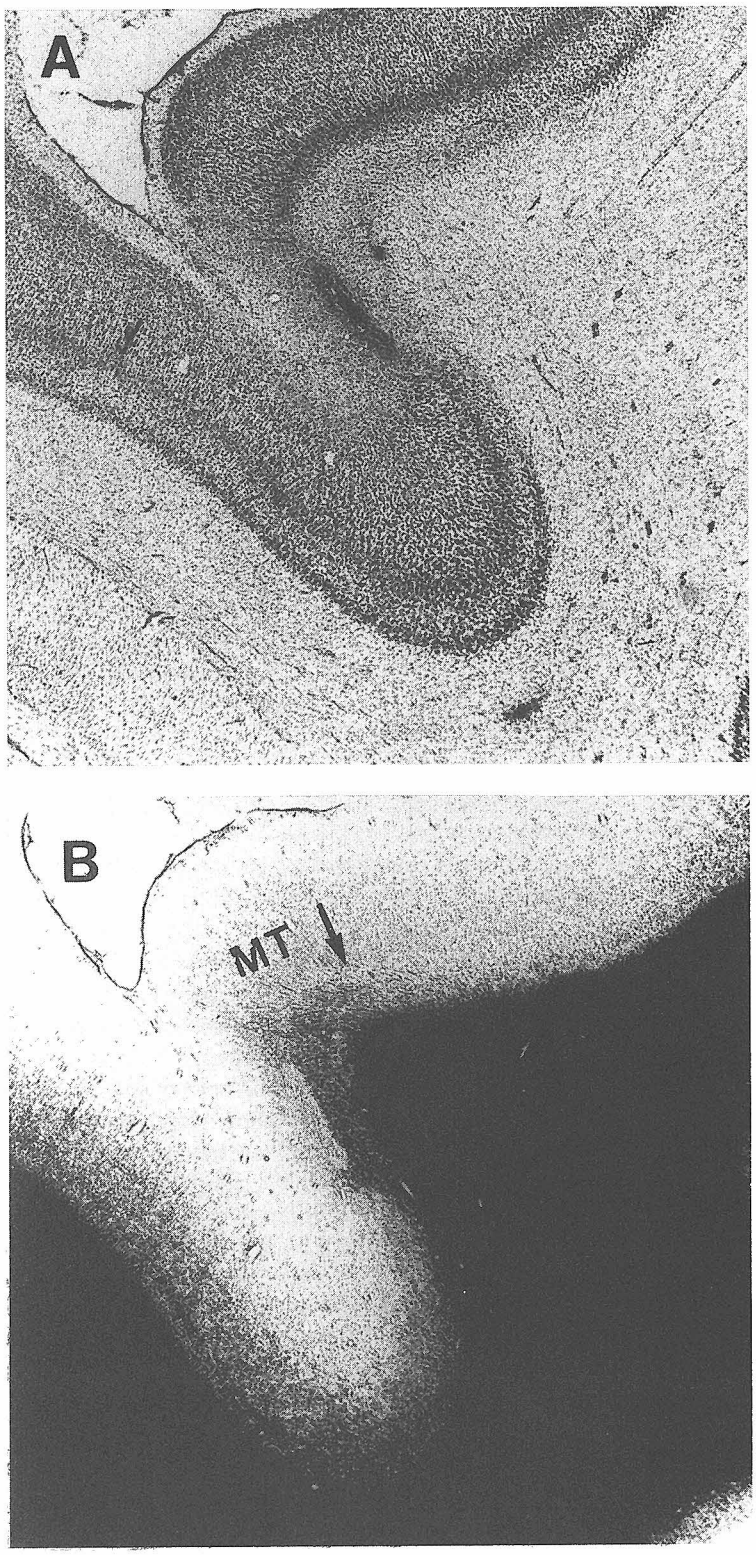

**Figure 11.** *Schematic of psychophysical motion tasks in lesion experiments.* The subject fixated on a small light-emitting diode, and test stimuli were presented at three locations in the upper, middle, and lower contralateral visual field. The location of appearance of the stimulus was randomized from trial to trial for the three locations. The test stimuli were from either the shear-motion detection task or the 3-D cylinder structure-from-motion task.

Figure 11 schematically shows the task the monkeys performed in the lesion experiments. They fixated on a small fixation point, and we measured shear-motion or structure-from-motion thresholds at three locations in the upper, middle, and lower contralateral visual fields. Control data were obtained over many days before the lesion, and an examination of prelesion data in Figure 13 shows that the thresholds measured from day to day were quite stable and reproducible.

Before each monkey was lesioned, we mapped the retinotopic organization of MT using single cell recording techniques. We then made a lesion at one of the three locations in the retinotopic map corresponding to the visual field locations from which psychophysical data had been obtained. We made the lesions in awake monkeys by injecting ibotenic acid through a microsyringe into MT. The use of ibotenic acid was important since it destroys cell bodies but does not damage fibers of passage (Figure 12). As a result, the optic radiation fibers passing below MT were not damaged, thus ensuring that the lesion was restricted to MT.

Figure 13 shows the effects on shear thresholds of an ibotenic acid lesion in the lower field representation of MT. The monkey was tested 1 hour after the lesion and already showed a dramatic increase in thresholds that was largely confined to the retinotopic locus of the lesion. The largest effect was seen 1 day after the lesion. Interestingly, the defect largely recovered in the next 2–3 days after that. This recovery pattern was seen in both hemispheres tested with MT lesions.

It was important to show that this effect was specific for motion and did not simply reflect the fact that the monkey was blind at the retinotopic locus of the


A
B
MT

lesion. In Figure 14, the controls, plotted at the left, show the contrast thresholds before and 2 days after lesioning, while at the right are the shear thresholds before and 1 day after lesioning. The contrast thresholds were measured at the same retinotopic location as the motion thresholds, using sinusoidal contrast gratings. Note that the psychometric shear-motion function shifted to the right after lesioning indicating increased thresholds, whereas the contrast sensitivity function was unaffected by lesioning. Thus the MT lesion selectively affected motion perception without affecting contrast perception.

In one hemisphere we also tested the effects of MT lesion on the ability to detect structure from motion. The task used the same 3-D revolving cylinder stimulus described earlier. Figure 15 shows the results for detection tested at transitions from no structure to 100% structure before and after lesioning. Two days before the MT lesions, the monkey was getting almost 90% correct on this relatively easy task. However, 23 days after the lesion and long after the shear-motion thresholds had recovered, it could not do the task for test stimuli near the retinotopic locus of the lesion. When this location was tested, the monkey had a hit rate near chance and a large number of early responses as it switched to a timing strategy in an attempt to maximize the amount of juice reward obtained. However, when we tested a site 6° away in the same half of the visual field, the monkey could still do the task, although not as well as prior to the lesion. These were the last data collected from this monkey. They suggest that structure-from-motion detection deficits may be longer lasting than simple shear-motion detection deficits, although additional experiments are required to reproduce this finding.

In one case we also tested shear-motion and structure-from-motion thresholds after lesioning MST. Interestingly, both the shear-motion and 3-D structure-from-motion thresholds were greater at all three locations tested in the contralateral visual field, and the thresholds for both recovered in 3–4 days, unlike the differential recovery periods seen after MT lesion. Further MST lesion studies are in progress.

## DISCUSSION OF THE MOTION DEFICITS

Lesions of MT produce a scotoma that prevents the detection of shearing motion at the locus in the visual field corresponding to the site of the lesion in its retinotopic representation. Remarkably, this deficit disappears in a matter of a few days.

Newsome et al. (1985) tested the effects of ibotenic acid lesions of MT on smooth pursuit eye movement, a behavior that requires motion analysis, and found a retinotopic deficit in estimating the speed of the target during the early, open-loop phase of tracking. This deficit also disappeared in a few days. These experiments did not, of course, establish that monkeys could not perceive motion after MT

---

**Figure 12.** *Example of an ibotenic acid lesion in MT. A:* Nissl-stained section showing the loss of cell bodies as a result of the injection. The elongated cut in the tissue was produced by the needle used to inject the neurotoxin. *B:* Adjacent section counterstained for myelin. Note that except for the damage resulting from the needle, the myelinated fibers in MT were unaffected by the ibotenic acid lesion. The arrow indicates the border of MT. As can be seen in this section, the location of MT can be determined by its dense myelination.

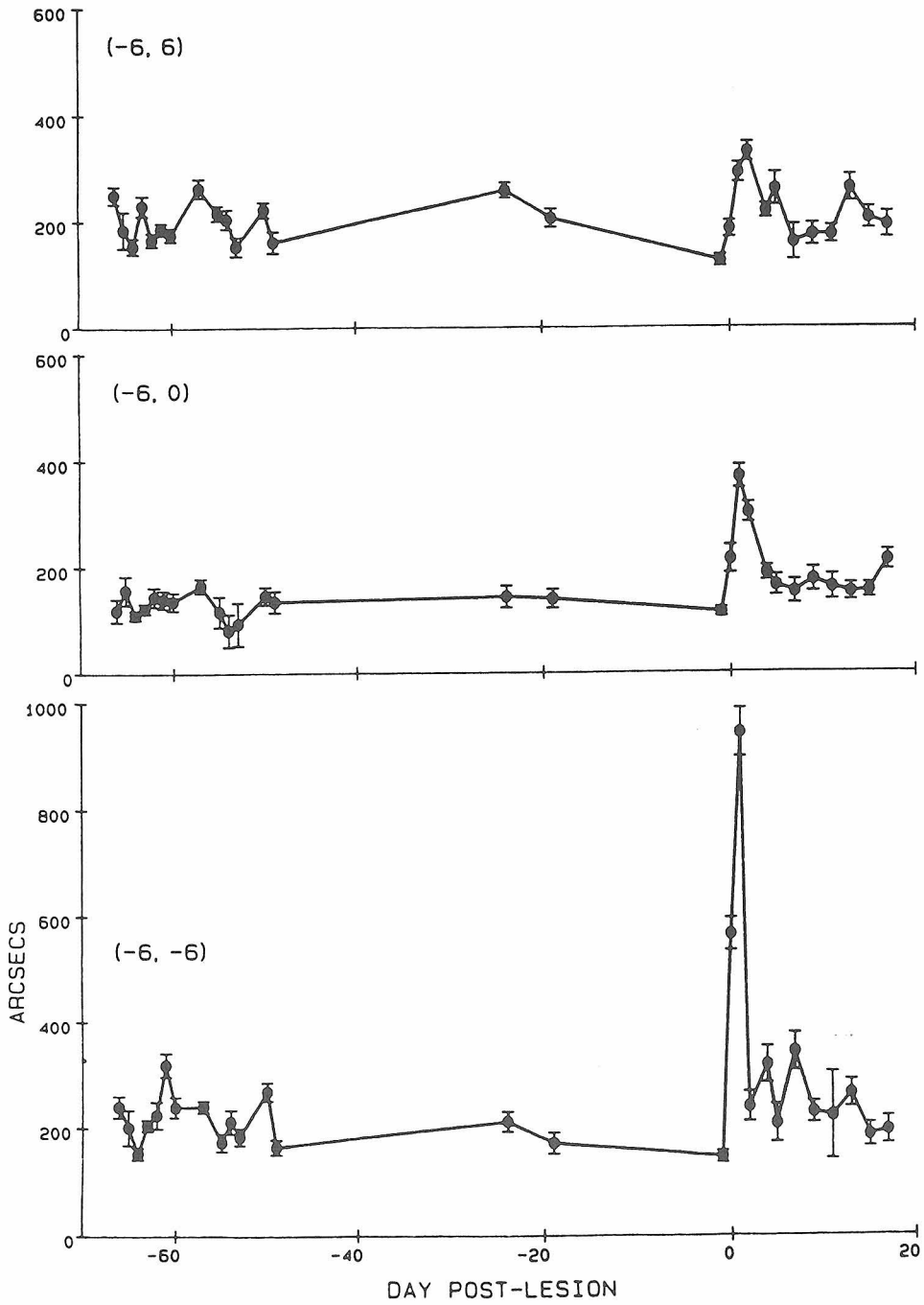



**Figure 13.** *The effect of ibotenic acid lesions in MT on shear-motion thresholds.* The 50% hit rates are plotted from psychometric curves in which the spatial and temporal frequencies are held constant and amplitude is varied. In this case the lesion was placed in the lower MT visual field representation in the contralateral (*right*) hemisphere. Note that the greatest effect was at the lower field location and that the thresholds increased after 1 day to more than 900 arc-sec. Also note that the thresholds recovered rapidly in a matter of a few days.

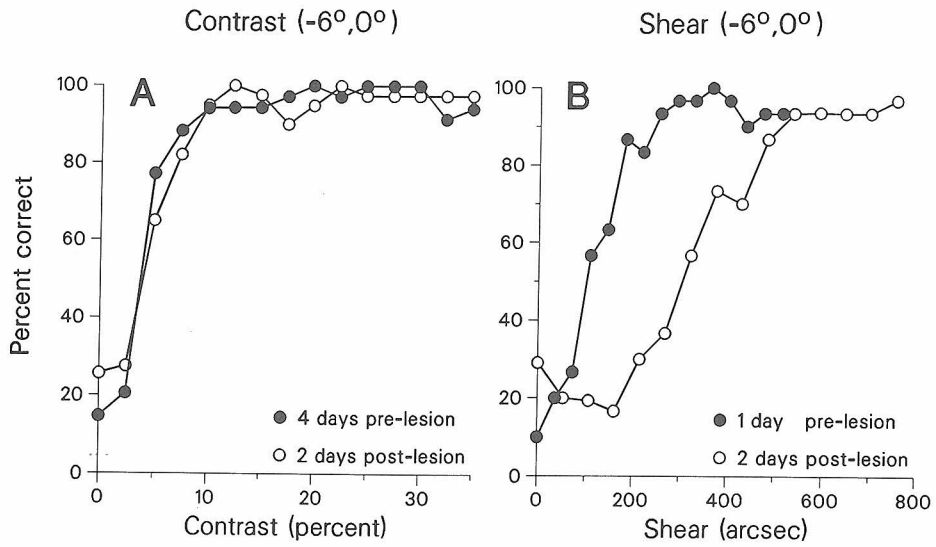



**Figure 14.** *A: Contrast thresholds measured 4 days before and 2 days after lesioning.* Contrast thresholds were measured by modulating the contrast of a sinusoidal spatial frequency grating. The subject performed the same reaction time task as the shear-motion detection task, except that instead of indicating the onset of motion he released the lever at the transition from no grating to a static contrast grating. *B:* Shear-motion thresholds recorded at the same retinal locus as the contrast data in *A* 1 day before and 2 days after lesioning. Note that there was a significant increase in motion threshold but no significant change in contrast threshold.

lesion, only that they could not use this information for smooth pursuit. However, our demonstration of a deficit in motion perception after MT lesion using similar techniques suggests that the pursuit deficit was a result of a motion perception deficit. Newsome and Paré (1986) reported a deficit in the ability of monkeys to determine the direction of correlated motion embedded in noise after MT lesions. Again, this deficit was transitory.

The experiments described above indicate that MT is an important part of the motion-processing apparatus in normal monkeys and that damage to this area leads to compensatory changes that result in recovery of function. This recovery might result from reorganization within MT or from parallel visual pathways around MT being strengthened or recruited. One way to test these two possibilities is to completely destroy MT. If the motion thresholds do not recover, the return of function after more restricted lesions are made is due to compensatory changes within MT; on the other hand, if the thresholds do recover, parallel pathways must be involved. Parallel pathway mechanisms offer several testable hypotheses, among them: (1) a pathway that is known to carry motion information such as the V1-to-V3 projection or the V1-to-V2-thick-stripes-to-V3 projection is strengthened; (2) a visual pathway normally not associated with motion analysis undergoes functional metamorphosis; or (3) changes occur at levels above MT that receive input from MT and other motion-processing pathways. MT lesion would upset the balance in these higher centers, and compensation might include changing the weights of inputs from the different motion pathway. Of course, this list does not exhaust all possibilities, and any combination of them may also be responsible for the compensatory changes.

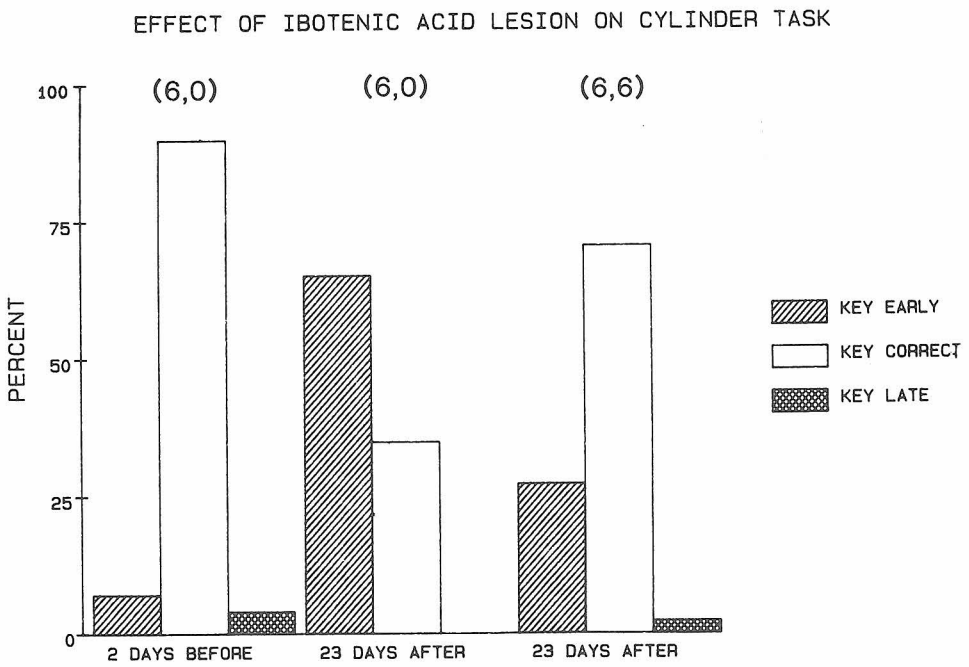



**Figure 15.** *Data for the reaction time task, in which a monkey detected the transition from noise to a completely structured revolving cylinder.* Two days before lesioning, the monkey got mostly hits at the retinal locus to be affected by the MT lesion (6,0); 23 days after the lesion, after shear detection thresholds recovered, it could not do the task at the retinal site affected by the lesion (6,0). It performed at a chance rate and had a large number of early responses. However, at a retinal locus tested away from the lesion-affected area (6,6) it could still do the task.

It is also important to consider the issue of training and motion experience on recovery; in the course of recovery, each animal performs several thousand trials on the motion detection task. If the animal were to be placed in a dark room for a week after the lesion rather than undergoing immediately rigorous testing, would the thresholds recover spontaneously, or would they remain high?

The one observation of a longer-lasting deficit in structure-from-motion perception after an MT lesion suggests either that MT is necessary for this computation and no other area can subserve its function or that MT provides essential preprocessing to other areas such as MST that might be involved in computation. Further experiments will have to be done to establish whether MT lesions result in permanent structure-from-motion deficits. It is interesting to consider the results of the one MST lesion we carried out. This lesion produced a shear motion detection deficit that was not as retinotopically restricted as that seen after MT lesion. These results are consistent with those of Wurtz (1986) and his colleagues, who found that MST lesions produced speed estimation deficits in the initial open-loop phase of smooth pursuit throughout the retinotopic visual field. They are also not unexpected, given the large receptive fields and cruder retinotopic organization of MST (Van Essen, 1985; Maunsell and Newsome, 1987). The large receptive fields may have some bearing on the fast recovery. Based on the evidence, one could hypothesize that area MST is involved in structure-from-motion processing and that it

receives some essential preprocessing from MT. However, due to the relative lack of retinotopic organization and the large receptive fields, a small, restricted lesion of MST would produce only an imbalance in processing structure from motion in this area, and this could easily be compensated for by changes in MST due to a redundancy of the structure-from-motion processing apparatus in this area. This hypothesis predicts a more permanent structure-from-motion deficit only after a large or possibly complete bilateral destruction of MST.

## CONCLUSIONS

The experiments outlined here show that monkeys, like humans, can perceive structure from motion. Moreover, monkeys tested to detect shear motion and various types of structure from motion always showed similar performance as humans tested with the same paradigms. These results suggest that the motion analysis system in the monkey is the same as the motion analysis system in humans and that what we learn about this system from lesion or recording experiments in monkeys is also likely to be true of humans.

We found that the ability to detect structure from motion is a global phenomenon that requires integrating information across the entire flow field. Reaction time data has provided evidence that monkeys and humans both use percepts of 3-D surfaces to detect the 3-D cylinder. Information integrated over both space and time has been found to construct neural representations of 3-D surfaces. These experiments suggest that the nervous system recovers surfaces from motion, and these observations should help to determine what algorithms the brain might use in the computation of structure from motion of rigid objects.

Our lesion experiments and those of others establish that MT is part of a motion pathway responsible for the perception of motion. Deficits after MT lesion are specific to motion and do not affect contrast sensitivity. The remarkable recovery of shear-motion detection thresholds and smooth pursuit capabilities after MT lesion indicates that the nervous system compensates for damage to MT. Because the anatomy and functional organization of the extrastriate pathways are now fairly well understood, it is possible to do experiments that will provide insight into the neural mechanisms for compensation of brain damage. Finally, the single observation of a longer-lasting deficit in detecting structure from motion after MT lesions, if repeated and found to be permanent, would indicate that MT is necessary to this type of motion analysis.

## ACKNOWLEDGMENTS

We thank Dr. M. Husain for participating in some of these experiments, C. Andersen for editorial assistance, and D. Duffy for typing the manuscript.

## REFERENCES

Allman, J. (1987) Variations in visual cortex organization in primates. In *Neurobiology of Neocortex*, P. Rakic and W. Singer, eds., pp. 29–40, Wiley, Chichester.

Allman, J., F. Miezen, and E. McGuinness (1985) Stimulus specific responses from beyond the classical receptive field: Neurophysiological mechanism for local–global comparisons in visual neurons. *Annu. Rev. Neurosci.* **8**:407–430.

Andersen, R. A., C. Asanuma, G. Essick, and R. M. Siegel (1990) Corticocortical connections of anatomically and physiologically defined subdivisions within the inferior parietal lobule. *J. Comp. Neurol.* **296:**65–113.

Frost, B. J., and K. Nakayama (1983) Single visual neurons code opposing motion independent of direction. *Science* **13:**744–745.

Gnadt, J. W., and R. A. Andersen (1988) Memory related motor planning activity in posterior parietal cortex of macaque. *Exp. Brain Res.* **70:**216–220.

Golomb, B., R. A. Andersen, K. Nakayama, D. I. A. MacLeod, and A. Wong (1985) Visual thresholds for shearing motion. *Vision Res.* **25:**813–820.

Grzywacz, N. M., E. C. Hildreth, V. K. Inada, and E. H. Adelson (1986) The temporal integration of 3-D structure from motion: A computational and psychophysical study. In *Brain Theory,* G. Palm and A. D. Aertsen, eds., Springer-Verlag, Heidelberg.

Inada, V., E. C. Hildreth, N. Grzywacz, and E. H. Adelson (1986) The perceptual buildup of 3-D structure from motion. *Invest. Opthalmol. Vis. Sci. (Suppl.)* **27:**142.

Maunsell, J. H. R., and W. T. Newsome (1987) Visual processing in monkey extrastriate cortex. *Annu. Rev. Neurosci.* **10:**363–401.

Nakayama, K. (1981) Differential motion hyperacuity under conditions of common image motion. *Vision Res.* **21:**1475–1482.

Nakayama, K., and C. W. Tyler (1981) Psychophysical isolation of movement sensitivity by removal of familiar position cues. *Vision Res.* **21:**427–433.

Nakayama, K., G. H. Silverman, D. I. A. MacLeod, and J. Mulligan (1985) Sensitivity to shearing and compressive motion in random dots. *Perception* **14:**225–238.

Newsome, W. T., and E. B. Paré (1986) MT lesions impair discrimination of direction in a stochastic motion display. *Soc. Neurosci. Abstr.* **12:**1183.

Newsome, W. T., R. H. Wurtz, M. R. Dürsteler, and A. Mikami (1985) Deficits in visual motion processing following ibotenic acid lesions of the middle temporal visual area of the macaque monkey. *J. Neurosci.* **5:**825–840.

Orban, G. A., W. Spileers, B. Gulyas, and P. O. Bishop (1986) Motion in depth selectivity of cortical cells revisited. *Soc. Neurosci. Abstr.* **12:**584.

Rogers, B., and M. Graham (1979) Motion parallax as an independent cue for depth perception. *Perception* **8:**125–134.

Rogers, B., and M. Graham (1982) Similarities between motion parallax and stereopsis in human depth perception. *Vision Res.* **22:**261–270.

Saito, H., M. Yukio, K. Tanaka, D. Hikosaka, Y. Fukada, and E. Iwai (1986) Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *J. Neurosci.* **6:**145–157.

Sakata, H., H. Shibutani, Y. Ito, and K. Tsurugai (1986) Parietal cortical neurons responding to rotary movement of visual stimulus in space. *Exp. Brain Res.* **61:**658–663.

Siegel, R. M., and R. A. Andersen (1986) Motion perceptual deficits following ibotenic acid lesions of the middle temporal area (MT) in the behaving rhesus monkey. *Soc. Neurosci. Abstr.* **12:**1183.

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

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

Ullman, S. (1979) *The Interpretation of Visual Motion,* MIT Press, Cambridge, Massachusetts.

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

Van Essen, D. C. (1985) Functional organization of primate and visual cortex. In *Cerebral Cortex,* Vol. 3, A. Peters and E. G. Jones, eds., pp. 259–329, Plenum, New York.

Wurtz, R. H. (1986) Probing visual processing with eye movements. *Soc. Neurosci. Abstr.* **12:**361.

Zeki, S. M. (1974) Functional organization of a visual area in the posterior bank of the superior temporal sulcus of the rhesus monkey. *J. Physiol. (Lond.)* **236:**549–573.