Communicated by Antonio Damasio

Structure from Motion with Impaired Local-Speed and Global Motion-Field Computations

Lucia M. Vaina
Intelligent Systems Laboratory, College of Engineering,
Boston University, Boston, MA 02215 USA,
and
Harvard–MIT Division of Health Sciences and Technology,
Massachusetts Institute of Technology, Cambridge, MA 02139 USA

Norberto M. Grzywacz
Center for Biological Information Processing,
Department of Brain and Cognitive Sciences,
Massachusetts Institute of Technology, Cambridge, MA 02139 USA

Marjorie LeMay
Brigham and Women's Hospital, Harvard Medical School,
Boston, MA 02115 USA

Humans can recover the three-dimensional structure of moving objects from their changing two-dimensional retinal image, in the absence of other cues to three-dimensional structure (Wallach and O'Connell 1953; Braunstein 1976). In this paper, we describe a patient, A.F., with bilateral lesions involving the visual cortex who is severely impaired on computing local-speed and global-motion fields, but who can recover structure from motion. The data suggest that although possibly useful, global-motion fields are not necessary for deriving structure from motion. We discuss these results from the perspective of theoretical models for this computation.

1 Introduction

The recovery of structure from motion is the ability of humans to perceive the three-dimensional structure of objects from motion cues. This ability is one of the building blocks of our perception of three dimensions and has been extensively studied both psychophysically (Wallach and O'Connell 1953; Johansson 1975; Braunstein 1976; Todd 1984; Braunstein and Shapiro 1987; Grzywacz et al. 1987; Ramachandran et al. 1988; Siegel and Andersen 1988; Dosher et al. 1989; Husain et al. 1989; Hildreth

Neural Computation 2, 420–435 (1990)

et al. 1990) and theoretically (Ullman 1979, 1984; Longuet–Higgins and Prazdny 1981; Bobick 1983; Bruss and Horn 1983; Koenderink and van Doorn 1986; Grzywacz and Hildreth 1987). Perhaps the most important lesson from these studies is that to recover structure from motion the brain assumes that moving objects are rigid (Ullman 1979) or quasirigid (Ullman 1984; Koenderink and van Doorn 1986; Grzywacz and Hildreth 1987).

What computations does the visual system use in order to obtain structure from motion?

The most natural candidate for such computations is the velocity field. Theoretically, one can usually obtain the structure of a rigid object if its velocity field is known (Longuet–Higgins and Prazdny 1981; Bruss and Horn 1983). This is possible, because under the rigidity assumption, the distance between the object's features is constant, imposing simple geometric constraints on the velocities. Psychophysical studies of two- and three-dimensional structures from brief motion observations (Johansson 1975; Dosher et al. 1989) support the use of velocity in the recovery of structure from motion. It has been shown, however, that instantaneous velocity measurements alone may not be sufficient, since it takes long observation periods to perceive structure from motion accurately (Grzywacz et al. 1987; Siegel and Andersen 1988; Husain et al. 1989; Hildreth et al. 1990).

Alternative forms of computations on the image might also be used to recover structure from motion. One such computation involves the combination of feature positions over time without explicit use of velocity information (Ullman 1979, 1984; Grzywacz and Hildreth 1987). This computation also works by exploiting rigidity: The constancy of the distance between an object's features geometrically constrains the structures consistent with a set of its views. Other computations that possibly underlie the recovery of structure from motion include the use of geometric characteristics of the image, such as luminance boundaries (Ramachandran et al. 1988) and axis of symmetry (Braunstein and Shapiro 1987), which allow the use of partial velocity information, such as motion direction (Bobick 1983).

An important insight into which of the above computations are necessary for the recovery of structure from motion may come from studying the performance of patients with focal lesions involving the visual cortex. Selective visual motion deficits in patients who can recover structure from motion would indicate which motion measurements are not necessary for this computation.

2 The Patient

In this paper, we present data from a patient whose performance on psychophysical motion tasks indicates that the recovery of structure from motion does not necessarily rely on the computation of velocity-field.[1] The patient, A.F., a 60-year-old left-handed man, was studied following an acute hemorrhagic infarct in the posterior right hemisphere. The only significant aspect of his clinical background is a history of untreated hypertension. A magnetic resonance imaging (MRI) study, performed 3 months after the stroke, demonstrated a large, new, hemorrhagic infarction in the posterior right hemisphere involving the region of the temporal–parietal–occipital junction. The MRI study also revealed an old, smaller, lesion in the left hemisphere involving the same anatomical areas as on the right hemisphere. Figure 1 shows A.F.'s visual fields, and the anatomical loci of his bilateral lesions both in axial MRI and on a schematic representation of the lateral view of a human brain. On neuroophthalmological examination, he had good letter acuity (20/30 in both eyes), normal contrast sensitivity at spatial frequencies ranging from 0.2 to 9 cycles/degrees, normal saccadic eye movements to static targets, but impaired saccades to moving targets, and the optokinetic nistagmus was absent in all directions. He also substituted saccades for smooth pursuits when he followed a smoothly moving target. This occurred for both motions to the right and to the left.

On the neuropsychological evaluation with the Wechsler Adult Intelligence Scale-Revised, he obtained an average Verbal IQ of 104 and a severely depressed Performance IQ of 68. Reading, writing, and oral calculations were not impaired. He showed no deficits on visual semantic tasks. Color (Farnsworth–Munsell 100-Hues test), form (Efron 1968), and texture discrimination (Julesz 1984) were excellent. His drawing was poor, disorganized, and lacked perspective. Stereopsis, tested with Julesz stereograms (Julesz 1971), was lost at all disparities tested, which distressed him enormously, since previously he enjoyed anaglyphs and geometric puzzles. Monocular depth, tested informally, was good.

We followed the patient on a regular basis for 18 months, and his performance on the psychophysical tasks and the results of the neuroophthalmological examinations remained consistent throughout the period.

---

[1] The stimuli were generated and presented, and responses collected and analyzed, using a Macintosh IIcx computer with an extended 8 bit video card. Stimuli were presented in the center of the Macintosh standard RGB monitor with a resolution of 640 × 480 pixels with frequency of the vertical retrace interrupt of 66.7 Hz. The contrast linearity of the display was measured and found to hold up to 98% contrast. In all the psychophysical tasks reported here, the display consisted of dynamic random dot patterns, which were presented for 2 sec in each trial. The viewing distance was 65 cm and the dot size was 1.8 × 1.8 arcmin. The background in the display was black and the dots were white. The room illumination was maintained at the low photopic level and the subjects viewed the display binocularly. Most of the normal subjects were naive observers.

3 Experiments

The patient showed deficits in local speed discrimination and perception of motion coherence.

The speed discrimination task (Vaina 1988, 1989; Vaina et al. 1988, 1989) is illustrated in Figure 2A. The display consisted of two rectangles (4×2.5 degrees^2), each containing 20 dots,[2] moving in randomly distributed directions. In each rectangle, for a given trial, the speed of all the dots was constant, forming the basis for a grouping into a global, coherent, speed field. This was a two-alternatives forced-choice task in which the subject had to determine in which rectangle the dots were moving faster. The faster speed, presented at random in either rectangle, was always 4.9 degrees/sec and the speed ratios ranged from 1.1 to 5.5. Data from A.F. (20 trials per speed ratio) and normal controls (26 subjects with 10 trials per speed ratio) is shown in Figure 2B. Although the control subjects performed well (above 75% correct) at the 1.47 speed ratio, A.F. failed on this task up to a ratio of 5.5.

He was also dramatically impaired (Fig. 2D) on a motion coherence task similar to that of Newsome and Paré (1988) (Fig. 2C). As in their task, each trial presented a dynamic random dot field with a given percentage, p, of dots moving in a prespecified direction. The display was cyclic with a period of about 50 msec. In each cycle, a new set of dots was presented for 16 msec, after which it was erased and a new set plotted. (There were 15 dots in a 100 degrees^2 square aperture.) The probability that any given dot in a frame would be displaced (9 arcmin) in the chosen direction was p. This means that the probability for a given dot to move in the same direction for $n + 1$ frames decreased as p^n. The task was a four-alternatives forced-choice in which the subject was asked to determine the direction of motion, which could be up, down, left, or right. A staircase procedure was used to obtain the threshold of motion signal (percentage coherence) required to reliably determine the direction of motion. The threshold was obtained separately for presentations in left and right visual fields. The performance of A.F. was significantly below that of the normal controls (z test; $p < 0.001$) for presentations in either visual fields (Fig. 2D).[3] Furthermore, inspection of details of

---

[2] Dot density did not appear to play a significant role in A.F.'s performance in speed-related tasks. First, A.F. obtained similar results on the same task as in Figure 2A, but with dot density 50% lower. Second, A.F. scored at chance on a task of boundary detection by relative speed, in which the dot density was very high (50% light and 50% dark — Vaina et al. 1990). This deficit cannot be explained by an impaired ability to detect boundaries. The patient scored in the normal range on a task of boundary discrimination by relative directions and in tasks where the random-dot background was static (Vaina et al. 1990).

[3] The subjects were instructed to maintain fixation on a mark placed 2 degrees to the right or left from the border of the aperture. Eye movements were informally controlled by the examiner, and those trials in which the subject failed to maintain fixation were discarded. However, we believe that controlled fixation would not modify

the data indicates that the probability that the performance was above chance for percentages of coherence equal or lower than 17% was less than 0.01. (This percentage of coherence was the largest value in the staircase procedure that was below threshold for both visual fields.) The highest degree of coherence set by the staircase procedure for which reversals occurred was 78% (even in the worst normal subjects, reversals did not occur beyond 10% coherence).

In spite of these deficits, A.F. performed well in a structure-from-motion task. Although in certain conditions his performance was not as good as that of normal subjects, it always remained well above chance. This was surprising, since our previous studies showed that patients with more extensive right parietal lesions failed this task (Vaina 1988, 1989; Vaina et al. 1988, 1989). They never saw a *single* object, but instead they

Figure 1: *Facing page.* Visual fields and four axial magnetic resonance images (MRI) of the patient's brain, and their localization on a schematic lateral view of a human brain. The left and right visual fields are illustrated in A and B, respectively. The patient has a congruous loss of the left inferior visual field bilaterally and a minimal loss in the upper visual field. (C,D,E,F) The relevant slices in axial view of T2 weighted MRI (TR 2000/TE80). [Although A.F. is left handed (Oldfield–Geshwind questionnaire score of −50) the MRI scans show the anterior portion of the right hemisphere and the posterior portion of the left hemisphere to be wider than their counterparts, which is the type of asymmetry seen in the majority of normal right-handed individuals — LeMay and Kido 1978.] Labeling: bl, body the lateral ventricle; h, hemosiderin; fl, frontal lobe; oh, occipital horn of the lateral ventricle; pl, parietal lobe; po, parietal–occipital sulcus; sf, Sylvian fissure; ssc, supercellar cistern; t, trigone of the lateral ventricles; t1, superior temporal gyrus; t2, medial temporal gyrus; ptof, parietotemporo–occipital fossa (see G). (C) The picture shows the temporal horns of the lateral ventricles and some patchy hyperintensities (labeled as ptof) at their margins, which appear to be more intense in the right hemisphere. (D) The picture shows local tissue loss in the right hemisphere at the site of the hemorrhagic stroke in the parietal lobe just medial to the occipital horn of the lateral ventricle and anterior to the parietal occipital sulcus. There is also some patchy hyperintensity in the temporal lobe along the lateral margin of the right occipital horn. Localized hyperintense areas are seen in the left temporal lobe adjacent to the lateral proximal margin of the occipital horn. Small areas of increased signal are scattered in the basal ganglia and in the deep white matter by the bodies of the lateral ventricles, which may be associated with small vessel ischemic changes. Images E and F show bilaterally larger patchy areas of hyperintensity posterior, medial, lateral, and above of the trigones and at the margins of the bodies of the lateral ventricles. Tortuous narrow bands of hemosiderin (labeled h) are seen in both images at the site of the recent hemorrhage. *Continued next page.*

---

the conclusions of this paper, because in an earlier similar test without fixation point, A.F. was still impaired.

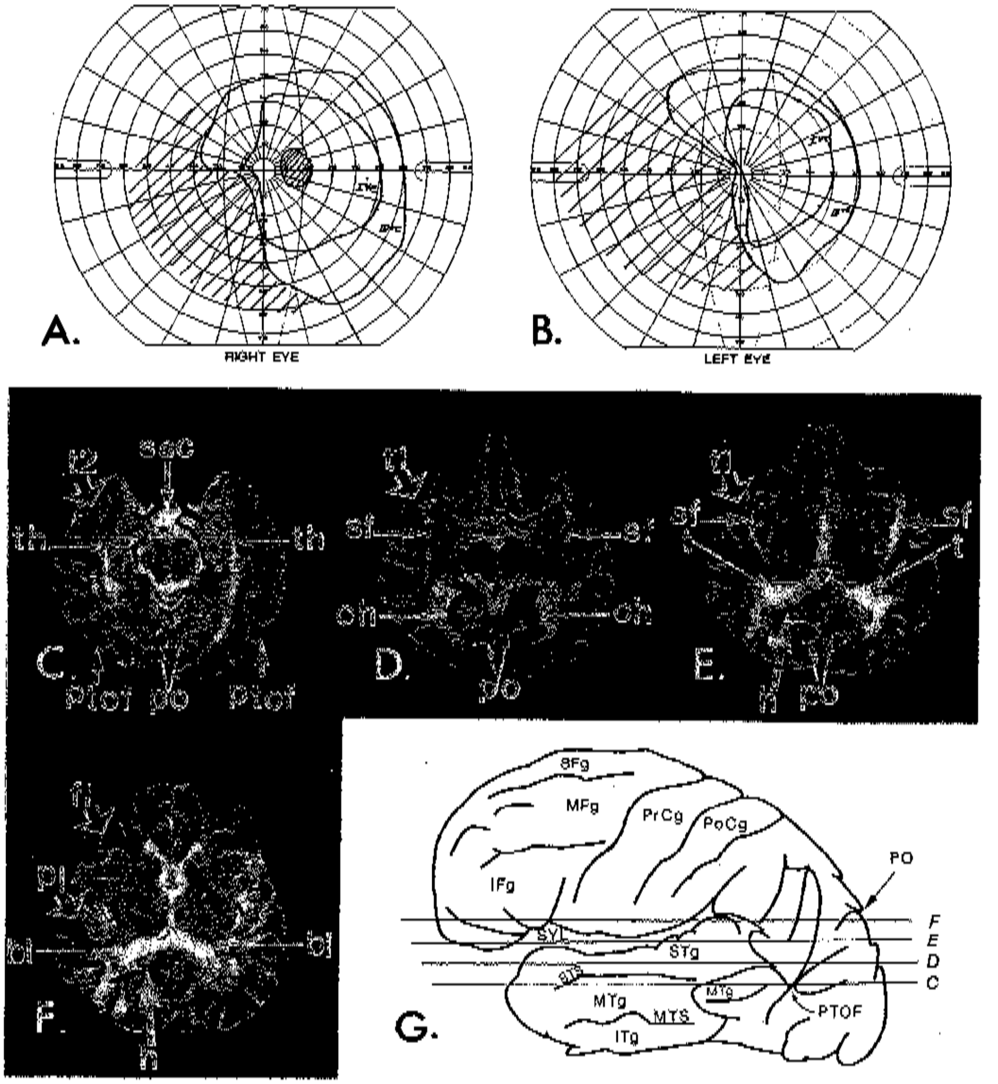


Figure 1: *Continued.* (G) A schematic drawing of the adult human brain showing the major gyri and sulci in lateral view. Labeling: IFg, inferior frontal gyrus; ITg, inferior temporal gyrus; MFg, middle frontal gyrus; MTg, middle temporal gyrus; PoCg, postcentral gyrus; PrCg, precentral gyrus; SFg, superior frontal gyrus; STg, superior temporal gyrus; STS, superior temporal sulcus; MTS, middle temporal sulcus; PO, parietal occipital sulcus; SYL, Sylvian fissure; PTOF, parietotemporo–occipital fossa (see Figure 271 of Polyak 1957); several authors suggested that it corresponds to the human homologue of the macaque, MT. The horizontal lines correspond to the levels of the axial scans C, D, E, and F.

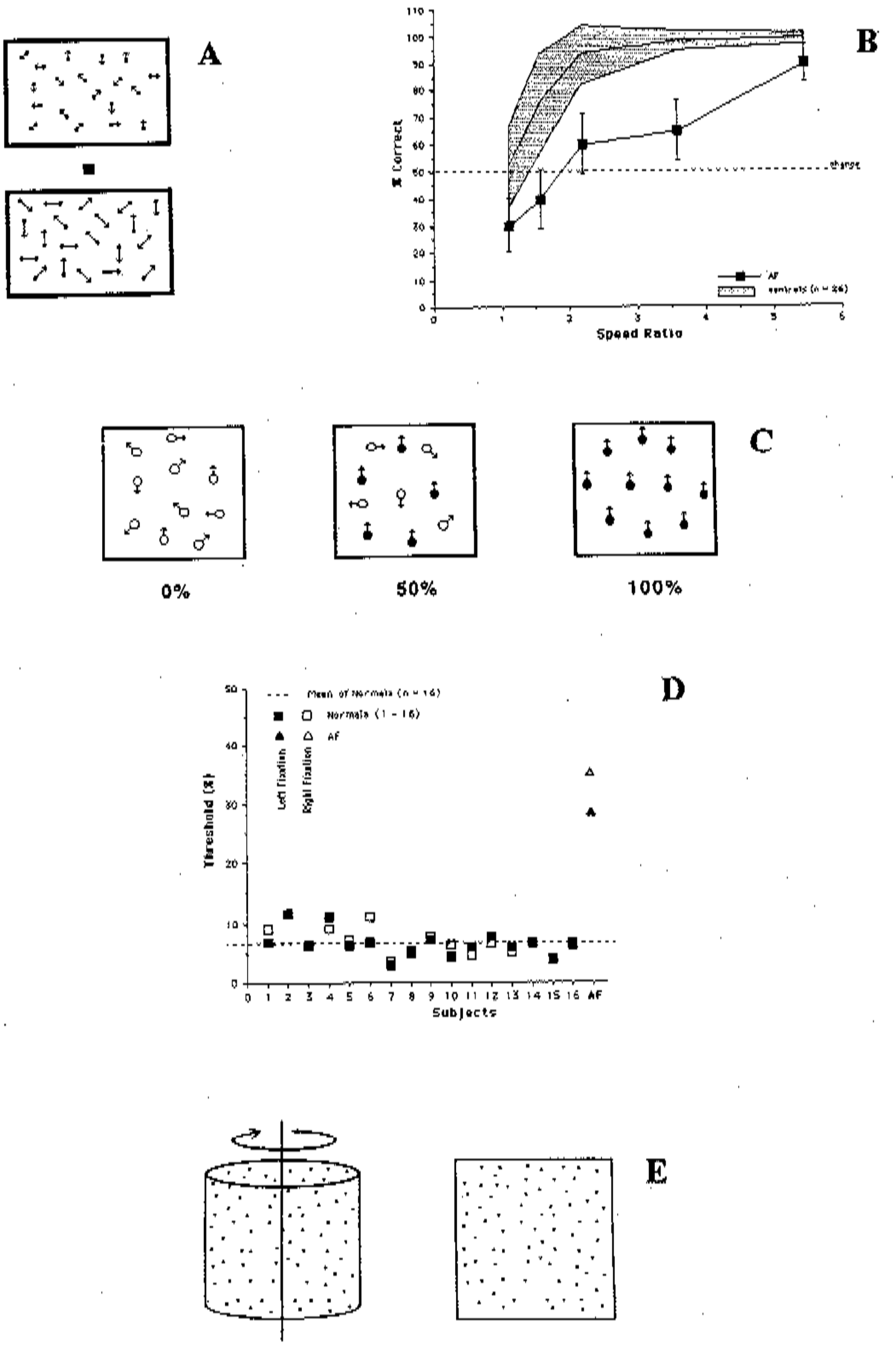
described the display as "birds flying," "ants crawling," "snow blown by the wind," or "just dots moving."

The display in this task showed two dynamic random-dots fields (3×3 degrees^2) each defined by 128 dots[4] with finite point lifetime. At the end of its lifetime, a point disappeared and was replotted in a new random location (within the boundary of each display) and began a new trajectory. One of the fields portraited the orthographic projection of a hollow cylinder rotating with 30 degrees/sec around its vertical axis. The average displacement of the dots between two frames was 5.6 arcmin. The rotation was simulated by 50 frames defined by dots lying at random locations on the cylinder's transparent surface (Figure 2E). The other random dot field contained an unstructured stimulus generated by randomly shuffling the velocity vectors present in the structured display and thereby destroying the spatial relationship between vectors. In both the structured and unstructured displays were 50 frames and the duration of each frame was 33 msec. The spatial positions (left and right) of the structured and unstructured field were randomly assigned. In one version of the task, the point lifetime was held constant at 400 msec and the amount of structure varied by shuffling an a priori determined percentage of the velocity vectors. In the second version of the task, the structure was held constant (100%) and the point lifetime of the dots in both the structured and unstructured displays varied. For both versions, there were 20 trials for each condition. We used a two-alternative forced-choice task in which the subject was asked to judge which of the two random dots fields represented a better cylinder.

Figure 2 (Facing page): Psychophysical motion tasks and data from A.F. and normal subjects. (A, C, E) Illustrate the paradigms used; for details of the experimental design see text. (A) A schematic representation of the stimuli employed in the local speed–discrimination task. (B) Results of the speed–discrimination task. The graph plots the percentage of correct answers as a function of the speed ratios between the two rectangles. The data for the normals are presented as a shaded area representing the mean ±1 standard deviation. A.F.'s data points present the means and the standard errors. For all ratios equal to or larger than 1.47, A.F.'s performance was significantly impaired compared to the normals. (C) A schematic representation of the stimuli employed in the motion-coherence task. (D) Results of the motion–coherence task. The graph plots the thresholds of percentage coherence required by normal subjects and A.F. to discriminate the net direction of motion. A.F.'s threshold is significantly elevated compared to the normals. (E) A schematic representation of the stimuli employed in the structure-from-motion task. *Continued on p. 428.*

---

[4] The results of normal subjects and A.F. were not statistically significantly different when the dot density was 50% lower.


A
B
110
100
90
80
70
60
50
40
30
20
10
0
% Correct
chance
AF
controls (n = 26)
0
1
2
3
4
5
6
Speed Ratio
C
0%
50%
100%
D
50
40
30
20
10
0
Threshold (%)
Mean of Normals (n = 16)
Normals (1 - 16)
AF
Left Fixation
Right Fixation
0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 AF
Subjects
E

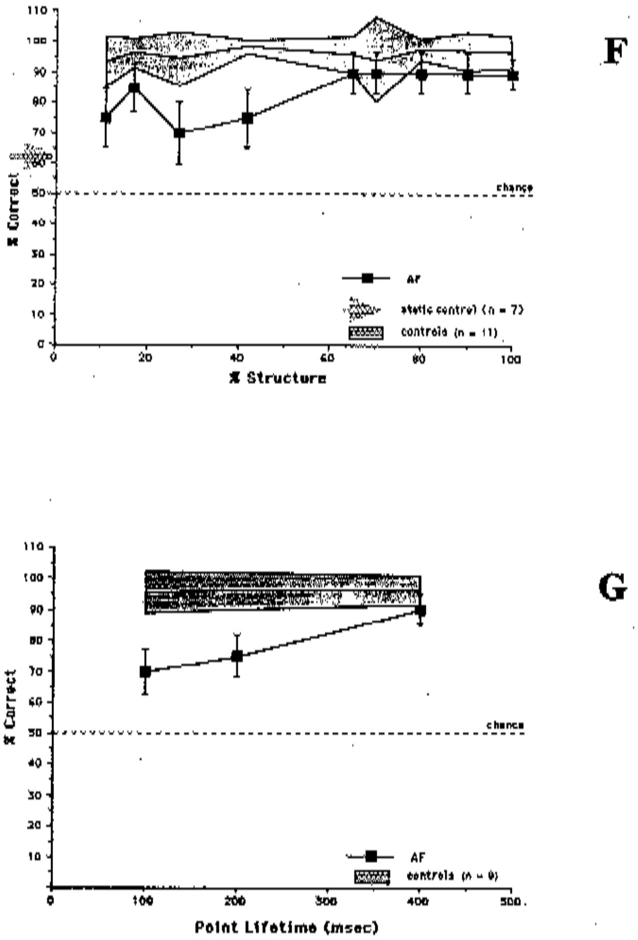


Figure 2: (F,G) Results of the structure–from–motion task. For plotting the data, we followed the same procedures as in B. (F) The graph plots the percentage of correct answers as a function of the fraction of structure in the rotating cylinder. The arrow represents the mean results for normal subjects with the task modified such that the frames in the cylinder were presented in random order. A.F.'s performance was in the normal range for percentages of structure in the cylinder equal or above 65%, but was significantly below that of the normal subjects in the trials where the structure was equal or below 42%. As indicated by the arrow, the subjects were probably not using texture cues to solve this task. (G) The graph plots the percentage of correct answers as a function of the point lifetime of the dots. A.F.'s performance was in the normal range for the lifetime equal to 400 msec, but was significantly below that of the normal subjects in the trials with lifetimes equal to 200 and 100 msec.

The data presented in Figure 2F and G show that A.F.'s performance on this task was well above chance for all conditions. Even in the presence of noise, when the percentage of structure was reduced (Fig. 2F), his performance was in the normal range for percentages above 42%. A.F.'s performance was significantly reduced for percentages of structure equal or lower than 42% (χ^2 test, $p < 0.007$). On these percentages of structure his scores were significantly lower than those of the worst normal subjects. Similarly, his performance was also dependent on the value of the point lifetime (Fig. 2G). Statistical analysis (Cochran–Mantel–Haenszel test, $p < 0.03$) indicates that A.F.'s performance decreases as the point lifetime decreases. In addition, although his performance was in the normal range for point lifetime of 400 msec, he was significantly impaired for 200 and 100 msec. That A.F. could perform the task, even though he had a severe deficit on the motion coherence task, suggests that the ability of computing global motion fields *is not necessary* for deriving structure from motion.[5] For example, at 11% structure he still performed well above chance (Fig. 2F), whereas in the Newsome and Paré's task (Fig. 2C), the probability that he could obtain the global motion field at a coherence of 17% or less was very small ($p < 0.01$). However, since his structure-from-motion performance decayed with reduced structure, it is likely that global motion fields are used if available. His dependence on the point lifetime indicates that he needs longer temporal integration than normal subjects for a good recovery of structure from motion. A possible explanation for this dependence is that A.F. requires a longer integration time to measure speed (McKee and Welch 1985), and that such measurement might be useful for computing structure from motion. The same explanation might account for his poor performance on the local-speed discrimination task, because in that task, the dot lifetimes were short (133 msec). It is conceivable that for longer dot lifetimes his ability to discriminate speed might have been better.

4 Discussion

It is particularly interesting that A.F. was selectively impaired on local-speed and motion-coherence discrimination tasks, but not on structure-from-motion tasks, because the anatomical evidence suggests that his lesions may involve the neural circuitry supporting the human homologue of the monkey's middle temporal area (MT). Physiological and behavioral studies indicate that in monkeys MT is critical for all these

---

[5] In structure-from-motion tasks (Siegel and Andersen 1988; Vaina 1988, 1989; Vaina et al. 1988, 1989; Husain et al. 1989) like in the one presented here, one might be suspicious that texture cues contribute to the perception of three dimensionality. To control for this possibility, we repeated the task with the frames of the display containing the cylinder presented in random order. If texture cues were available, then the subjects should have been able to use them in making their choice. As Figure 2F shows, the normal subjects performed at chance under this condition.

computations (Zeki 1969; Maunsell and van Essen 1983). The conjecture that A.F.'s lesions disconnect bilaterally the human homologue of MT is supported by three sets of arguments. First, studies of changes in cerebral blood flow as monitored with positron emission tomography (PET) in a human observing low contrast moving stimuli and fast flicker (Miezin et al. 1987), and dynamic random-dot patterns (Lueck et al. 1989) showed significantly increased activity in the region of the fundus of the temporal–occipital–parietal fossa PTOF (Polyak 1957). This region was proposed as the possible correspondent of MT in humans (Miezin et al. 1987; Allman 1988; Lueck 1989; Zeki 1990). (Although there are limitations in the resolution of the PET equipment as compared to MRI, the former can localize the center of mass of the activity distribution well and, thus, is adequate for localizing MT) (Allman, personal communication.) Second, comparison of the myeloarchitecture pattern of human and monkey brains indicates that Flechsig area 16 (Flechsig 1920), which is the most myelinated area in the occipital–parietal cortex, might correspond to the heavily myelinated MT in the monkey (Allman 1977; Thurston et al. 1988). It is believed that Flechsig area 16 and PTOF correspond anatomically. As discussed in Figure 1, A.F.'s lesions appear to have disrupted the pathways to and from PTOF. Third, as discussed above, A.F. lacked smooth pursuits eye movements and it has been shown that MT is involved in the control of smooth pursuits (Komatsu and Wurtz 1988; Thurston et al. 1988).

From the putative disconnection of the human homologue of MT in A.F.'s lesions and his good performance on structure-from-motion tasks, we conjecture that MT is not necessary for structure from motion. This hypothesis seems to be at odds with Siegel and Andersen's report (1986) that lesions of MT with ibotenic acid apparently disrupt the recovery of structure from motion in monkeys. (It is possible, however, that A.F. might have recovered his ability to derive structure from motion, since we first gave him the task 3 months after the stroke.) There is no evidence, however, that in Siegel and Andersen's experiment the monkey performed a structure-from-motion task. The animal was required to detect a change from an unstructured to a structured motion velocity field, which portrayed a rotating surface of a cylinder covered with short-lived dots. The monkey might have responded, for example, to the ratio between the speeds at the center and corners of the display, which is higher for the cylinder than for the unstructured movie. Or alternatively, the monkey could have just noticed the increase in speed of dots at the center of the display when it became structured. Although these tricks were available to A.F. as well, we have direct evidence that he could recover structure from motion: Prior to the structure-from-motion task described here, we gave him a control display, which consisted of a dynamic random dot field portraying a rotating cylinder and asked him to describe verbally what he saw. The only instruction we gave him was that dots portrayed a moving object. Without any hesitation, A.F. reported seeing a rotating cylinder.

What types of computations may underlie A.F.'s ability to recover structure from motion? Four possibilities come to mind. First, we have seen that complete deterioration of the ability to compute global-motion fields for low fraction of coherence does not imply the destruction of the ability to obtain structure from motion under similar noise conditions. However, the presence of noise in the structure-from-motion task appears to have produced a deterioration of A.F.'s performance (but his scores remain consistently above chance). Thus, though *not necessarily*, the computation of global-motion fields may help to recover structure from motion. Second, as suggested by several computational studies, local-speed measurements may also underlie this computation. The patient's poor performance on the local-speed discrimination task rules out theories that require precise speed measurements (Longuet–Higgins and Prazdny 1981; Bruss and Horn 1983). However, rough speed estimates may help the perception of three-dimensionality, since there is a large speed ratio between the center and edges of the cylinder in our structure-from-motion task. Such cue was available to A.F., as this ratio is larger in the cylinder than is his speed-discrimination threshold (Fig. 2B). As we speculated above, the partial deterioration of his performance with reduced point lifetime is consistent with local-speed cues being used to recover three-dimensional structure from motion. Third, another strategy for this computation is combining the positions of image features over time, which may not require any explicit instantaneous motion measurement (Ullman 1979, 1984; Grzywacz and Hildreth 1987). However, all position-based strategies that have been proposed so far would break down for finite point lifetime experiments (Todd 1984; Siegel and Andersen 1988; Husain et al. 1989). This implies that A.F. did not use any such strategy to perceive the rotating cylinder. A recent algorithm that combines a position-based strategy with surface interpolation was shown to perform well under finite point lifetime conditions (Hildreth and Ando, personal communication), and it may provide a plausible strategy. Fourth, several studies speculated that specific geometric characteristics of the image, such as boundaries (Ramachandran et al. 1988) or axes of symmetry (Braunstein and Shapiro 1987), may contribute to succesful structure from motion processes. For example, a computational study showed that if in a rotating object the axis of rotation is known, then measuring the direction of motion and the projected distance from the axis are sufficient for recovering the structure of the object (Bobick 1983). The information necessary to apply this strategy might have been available to A.F. The axis of rotation was easy to locate in our structure-from-motion task, because the axis was always in the middle of the display. Also, the only two directions of motion in the display were 180 degrees apart. The discrimination of these directions was probably an easy task for A.F., since indirect evidence indicates that A.F. could discriminate perfectly motion directions that differed by 37 degrees or more. This evidence comes from his performance on Hildreth's task of boundary-localization on the basis of direction of motion (Hildreth 1984).

The patient described in this paper has rather specific motion deficits (for another example of a motion-impaired patient see Zihl et al. 1983 and Hess et al. 1989). In particular, he is severely impaired on the discrimination and use of local-speed and on the perception of coherent motion. Thus, his good performance on structure-from-motion tasks is intriguing, since those are fundamental motion measurements, which furthermore are *exclusively* used by many computational theories of structure from motion.

5 Acknowledgments

We are indebted to Ellen Hildreth, Bill Newsome, and Richard Andersen for critically reading this paper. Also, thanks are due to Thomas Kemper for useful comments on the neuroanatomical findings. We thank the Neuropsychology and the Young Stroke Units of the New England Rehabilitation Hospital for referring the patient to us and for allowing us to use the results of the neuropsychological and clinical examinations. We also thank Don Bienfang for performing the neuroophthalmological evaluation. Al Choi programmed the software for the psychophysical tasks. L. M. V. and M. L. were supported by NEI Grant EY07861-01 and by the Boston University Seed Grants for Biomedical Research, 7712-5-ENG-909. N. M. G. was supported by NSF Grant BNS-8809528, by the Sloan Foundation, by a grant to Tomaso Poggio and Ellen Hildreth from the Office of Naval Research, Cognitive and Neural Systems Division, and by a grant to Tomaso Poggio, Ellen Hildreth, and Edward Adelson from the NSF.

References

Allman, J. M. 1977. Evolution of the visual system in early primates. In *Progress in Psychology, Physiology, and Psychiatry*, J. Sprague and A. N. Epstein, eds., pp. 1–53. Academic Press, New York.

Allman, J. M. 1988. The search for area MT in the human visual cortex. In *Proceedings of the European Brain and Behavior Winter Conference: Segregation of Form and Motion, Tübingen, West Germany*, J. Rauschaker and G. Moon, eds., University of Tübingen, Germany.

Bobick, A. 1983. A hybrid approach to structure from motion. In *Proceedings of the ACM Interdisciplinary Workshop Motion: Representation and Perception*, pp. 91–109. Association for Computing Machinery, New York.

Braunstein, M. L. 1976. *Depth Perception Through Motion.* Academic Press, New York.

Braunstein, M. L., and Shapiro, L. R. 1987. Detection of rigid motion in fixed-axis and variable-axis rotations. *Invest. Ophthalmol. Vis. Sci.* **28**, 300.

Bruss, A., and Horn, B. K. P. 1983. Passive navigation. *Comput. Vision Graph. Image Proc.* **21**, 3–20.

Dosher, B. A., Landy, M. S., and Sperling, G. 1989. Kinetic depth effect and optic flow — I. 3D shape from Fourier motion. *Vision Res.* **29**, 1789–1813.

Efron, R. 1968. What is perception? In *Boston Studies in the Philosophy of Science*, Vol. 4, R. S. Cohen and M. W. Wartofsky, eds., pp. 137–173. D. Reidel, Dordrecht, The Netherlands.

Flechsig, P. 1920. *Anatomie des Menschlichen Gehirns und Rückenmarks auf Myelogenetischer Grundlage.* G. Thieme, Leipzig, Germany.

Grzywacz, N. M., and Hildreth, E. C. 1987. The incremental rigidity scheme for recovering structure from motion: Position vs. velocity based formulations. *J. Opt. Soc. Am., A* **4**, 503–518.

Grzywacz, N. M., Hildreth, E. C., Inada, V. K., and Adelson, E. H. 1987. The temporal integration of 3-D structure from motion: A computational and psychophysical study. In *Organization of Neural Networks*, W. von Seelen, W. G. Shaw, and U. M. Leinhos, eds., pp. 239–259. VCH Publishers, Weinhein, FRG.

Hess, R. H., Baker, C. L., and Zihl, J. 1989. The "motion-blind" patient: Low-level spatial and temporal filters. *J. Neurosci.* **9**, 1628–1640.

Hildreth, E. C. 1984. *The Measurement of Visual Motion.* MIT Press, Cambridge, MA.

Hildreth, E. C., Grzywacz, N. M., Inada, V. K., and Adelson, E. H. 1990. *Percept. Psychophys.*, in press.

Husain, M., Treue, S., and Andersen, R. A. 1989. Surface interpolation in three–dimensional structure from motion. *Neural Comp.* **1**, 324–333.

Johansson, G. 1975. Visual motion perception. *Sci. Am.* **232**, 76–88.

Julesz, B. 1971. *Foundation of Cyclopean Perception.* University of Chicago Press, Chicago.

Julesz, B. 1984. Toward an axiomatic theory of preattentive vision. In *Dynamic Aspects of Neocortical Functions*, G. M. Edelman, W. E. Gall, and W. M. Cowan, eds., pp. 585–612. John Wiley, New York.

Koenderink, J. J., and van Doorn, A. J. 1986. Depth and shape from differential perspective in the presence of bending deformations. *J. Opt. Soc. Am., A* **3**, 242–249.

Komatsu, H., and Wurtz, R. H. 1988. Relation of cortical areas MT and MST to pursuit eye movements. I. Localization and visual properties of neurons. *J. Neurophysiol.* **60**, 580–603.

LeMay, M., and Kido, D. K. 1978. Asymmetries of the cerebral hemispheres on computed tomograms. *J. Comput. Assist. Tomogr.* **2**, 471–476.

Longuet–Higgins, H. C., and Prazdny, K. 1981. The interpretation of moving retinal images. *Proc. R. Soc., Lond., B* **208**, 385–397.

Lueck, C. J., Zeki, S., Friston, K. J., Deiber, M. P., Cope, P., Cunningham, V. J., Lammertsma, A. A., Kennard, C., and Frackowiak, R. C. J. 1989. The colour center in the cerebral cortex of man. *Nature (London)* **340**, 386–389.

Maunsell, J. H. R., and 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. *J. Neurophysiol.* **49**, 1127–1147.

McKee, S. P., and Welch, L. 1985. Sequential recruitment in the discrimination of velocity. *J. Opt. Soc. Am., A* 2, 243–251.

Miezin, F. M., Fox, P. T., Raichle, M. E., and Allman, J. M. 1987. Localized responses to low contrast moving random dot patterns in human visual cortex monitored with positron emission tomography. *Neurosci. Abst.* **13**, 631.

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

Polyak, S. 1957. *The Vertebrate Visual System.* University of Chicago Press, Chicago.

Ramachandran, V. S., Cobb, S., and Rogers–Ramachandran, D. 1988. Perception of 3-D structure from motion: The role of velocity gradients and segmentation boundaries. *Percept. Psychophys.* **44**, 390–393.

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

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

Thurston, S. E., Leigh, R. J., Crawford, T., Thompson, A., and Kennard, C. 1988. Two distinct deficits of visual tracking caused by unilateral lesions of cerebral cortex in humans. *Ann. Neurol.* **23**, 266.

Todd, J. T. 1984. The perception of three–dimensional structure from rigid and nonrigid motion. *Percept. Psychophys.* **36**, 97–103.

Ullman, S. 1979. *The Interpretation of Visual Motion.* MIT Press, Cambridge, MA.

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

Vaina, L. M. 1988. Effects of right parietal lobe lesions on visual motion analysis in humans. *Invest. Ophthalmol. Vis. Sci.* **29**, 434.

Vaina, L. M. 1989. Selective impairment of visual motion interpretation following lesions of the right occipital parietal area in humans. *Biol. Cybernet.* **61**, 347–359.

Vaina, L. M., LeMay, M., Naili, S., Amarillio, P., Bienfang, D., Montgomery, C., and Thomazeau, Y. 1988. Deficits of visual motion analysis after posterior right hemisphere lesions. *Neurosci. Abst.* **14**, 458.

Vaina, L. M., LeMay, M., Choi, A., Kemper, T., and Bienfang, D. 1989. Visual motion analysis with impaired speed perception: Psychophysical and anatomical studies in humans. *Neurosci. Abst.* **15**, 1256.

Vaina, L. M., LeMay, M., Bienfang, D. C., Choi, A. Y., and Nakayama, K. 1990. Intact "biological motion" and "structure from motion" perception in a patient with impaired motion mechanisms. *Vis. Neurosci.*, in press.

Wallach, H., and O'Connell, D. N. 1953. The kinetic depth effect. *J. Exp. Psychol.* **45**, 205–217.

Zeki, S. M. 1969. Representation of central visual fields in prestriate cortex of monkey. *Brain Res.* **14**, 271–291.

Zeki, S. M. 1990. The form vision of achromatopsic patients. In *"The Brain," Abstracts of the Cold Spring Harbor Symposium on Quantitative Biology*, p. 16.

Zihl, J., von Cramon, D., and Mai, N. 1983. Selective disturbance of movement vision after bilateral brain damage. *Brain* **106**, 313–340.

Received 16 March 90; accepted 30 July 90.