



Estimating Heading During Eye Movements

CONSTANCE S. ROYDEN,*† JAMES A. CROWELL,* MARTIN S. BANKS*

Received 30 November 1993; in revised form 28 February 1994

In eight experiments, we examined the ability to judge heading during tracking eye movements. To assess the use of retinal-image and extra-retinal information in this task, we compared heading judgments with executed as opposed to simulated eye movements. In general, judgments were much more accurate during executed eye movements. Observers in the simulated eye movement condition misperceived their self-motion as curvilinear translation rather than the linear translation plus eye rotation that was simulated. There were some experimental conditions in which observers could judge heading reasonably accurately during simulated eye movements; these included conditions in which eye movement velocities were 1 deg/sec or less and conditions which made available a horizon cue that exists for locomotion parallel to a ground plane with a visible horizon. Overall, our results imply that extra-retinal, eye-velocity signals are used in determining heading under many, perhaps most, viewing conditions.

Optic flow Motion Pursuit eye movement Heading

INTRODUCTION

During self-motion, we must determine the direction of our movement with respect to the body in order to reach a desired destination while avoiding obstacles. Our ability to do so is remarkable considering that images move on the retina, the eyes move in the head, the head moves on the body, and objects can move within the scene. Consider, for example, an outfielder running to catch a fly ball: while moving the eyes and head to track the ball, he must determine the future position of the moving ball with respect to the future position of his body in order to place the hand in the correct position at the right time (Chapman, 1968; McLeod & Denes, 1993).

How does the visual system process the available information to achieve an accurate estimate of heading relative to the head or body? Under ordinary circumstances, the nervous system has information from many sources including visual, vestibular, oculomotor, and motor systems, that could aid the estimation. However, flight simulators and motion pictures create a strong sensation of a direction of self-motion, which suggests that visual stimulation alone can provide much of the

information needed to estimate heading. We know little, however, about the accuracy of such estimates, particularly when the simulated aircraft or camera rotates. In this paper, we examine the ability to estimate heading in human observers when rotation occurs due to eye movements.

To describe the difficulties that arise in estimating heading during an eye movement, it is useful to explain some properties of the retinal image motions that accompany observer motion through a rigid scene. Figure 1(A) depicts the pattern of retinal image motions created by forward translation (linear motion) across a ground plane while holding the eyes and head in fixed position. Flow is directed away from the so-called *focus of expansion* (FOE), which corresponds with the direction of translation or heading. Gibson (1950, 1966) hypothesized that people determine heading by locating this focus of expansion.

The pattern of retinal image motions becomes more complex when the observer translates while maintaining fixation on a point off to the side. Figure 1(B) portrays the flow field resulting from the same translation as in Fig. 1(A) while the observer makes an eye movement. This motion does not produce a focus of expansion in the image corresponding to the heading; the closest approximation to a focus corresponds to the point of fixation (Regan & Beverley, 1982). Consequently, the direction of self-motion cannot be determined by locating a focus (or singularity) in the retinal flow field. This presents a serious problem for many older models (Clocksin, 1980; Gibson, Olum & Rosenblatt, 1955; Gordon, 1965; Lee, 1980) because people in everyday settings frequently make eye and head movements while walking, running and driving (Cutting, 1986).‡ The same

*School of Optometry and Department of Psychology, University of California—Berkeley, Berkeley, CA 94720, U.S.A. [Email marty@john.berkeley.edu].

†Present address: Department of Computer Science, Wellesley College, Wellesley, MA 02181, U.S.A.

‡Gibson appreciated this point and attempted to find a computational solution to it (see, e.g. Gibson *et al.*, 1955). Although he never succeeded, his insight that flow due to translation has different properties than flow due to rotation led to later computational solutions.

point holds for automated systems with rotational components in the motion of the platform and/or camera (Carne, 1992).

The problems introduced by observer rotations can best be appreciated by considering the flow equations (1) for an observer moving through a rigid environment. At any instant, the observer's motion can be described as the sum of a translation (T_x, T_y, T_z) and a rotation (R_x, R_y, R_z). A point P in the scene with coordinates (X, Y, Z) projects onto the retina, represented by an image plane behind the origin, at coordinates (x, y) where $x = -X/Z$ and $y = -Y/Z$. Differentiating these equations with respect to time and introducing the translation and rotation components (Longuet-Higgins & Prazdny, 1980) yields:

$$\begin{aligned} v_x &= (xT_z + T_x)/Z - xyR_x + (1 + x^2)R_y + yR_z \\ v_y &= (yT_z + T_y)/Z - (1 + y^2)R_x + xyR_y - xR_z \end{aligned} \quad (1)$$

where v_x and v_y are the horizontal and vertical velocity components of the image of P on the image plane. These equations allow one to calculate the retinal flow field for various scenes and observer motions. The velocities v_x and v_y can be further separated into translational and rotational components:

$$\begin{aligned} t_x &= (xT_z + T_x)/Z \\ t_y &= (yT_z + T_y)/Z \\ r_x &= -xyR_x + (1 + x^2)R_y + yR_z \\ r_y &= -(1 + y^2)R_x + xyR_y - xR_z. \end{aligned} \quad (2)$$

$$(3)$$

The translational components are zero at the retinal projection of the observer's heading, so setting t_x and t_y to zero yields the retinal coordinates of the heading $(t_x^*, t_y^*) = (T_x/T_z, T_y/T_z)$. Thus, the translational components contain all the information necessary for determining heading. It is, however, unclear how to determine the translational components, t_x and t_y , in the presence of rotations because the retinal image data consist of velocity components, v_x and v_y , that are also affected by the rotational components, r_x and r_y ,

†The means of representing optic flow presented by Koenderink and van Doorn (1976) is not considered in this paper for the following reasons. Koenderink and van Doorn showed how optic flow can be described by four components: translation and three differential invariants, divergence, curl, and deformation. Several authors (e.g. De Bruyn & Orban, 1990; Regan & Beverley, 1982; Saito, Yukie, Tanaka, Hikosaka, Fukada & Iwai, 1986; Duffy & Wurtz, 1991) have argued that the maximum of divergence (div) frequently corresponds to the direction of self-motion. However, it generally does not because the correspondence between max of div and direction of self-motion is highly dependent on the geometry of the scene (Regan & Beverley, 1982). Koenderink and van Doorn (1981) showed how div could be used for estimating heading relative to a plane, but they never solved the general inverse mapping problem (going from flow field properties to observer motion and scene properties for arbitrary viewing conditions), so it is not clear that the components of their analysis could be used to deduce heading in general.

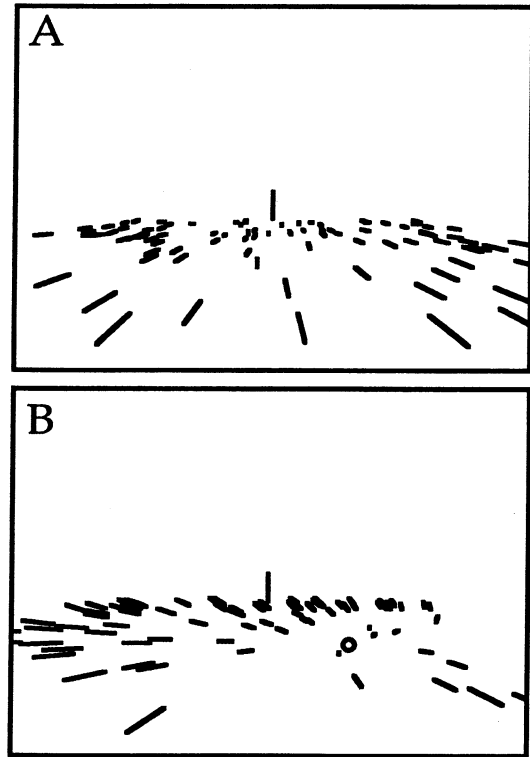


FIGURE 1. The pattern of retinal image motions associated with observer translation along a ground plane. (A) Observer is translating toward the vertical line while holding the eyes and head fixed. (B) Observer is translating toward the vertical line while making an eye movement to maintain fixation on the small circle.

[see equation (1)]. Therefore, one must somehow eliminate the confounding effects of the rotational components in order to compute the heading.

Recent theoretical efforts have focused on this problem. There are now numerous models falling into two categories that differ in their use of extra-retinal information. *Retinal-image models* (Longuet-Higgins & Prazdny, 1980; Longuet-Higgins, 1981; Regan & Beverley, 1982; Bruss & Horn, 1983; Rieger & Lawton, 1985; Waxman & Ullman, 1985; Droulez & Cornilleau-Peres, 1990; Heeger & Jepson, 1992; Hildreth, 1992; Perrone, 1992) employ retinal-image information only whereas *extra-retinal models* (von Holst, 1954) make use of eye-velocity information from extra-retinal sources.†

Although they use a variety of algorithms to estimate heading in the presence of rotational flow, all retinal-image models rely on the fact that flows due to translation and rotation have different properties. For instance, flow due to translation is depth-dependent while flow due to rotation is not, i.e. Z enters into equations (2), but not equations (3). Figure 2 schematizes one version of the retinal-image model (Longuet-Higgins & Prazdny, 1980) that first estimates the rotational flow components, then subtracts them from the original flow field, and finally estimates heading from the remainder.

Retinal-image models offer simplicity because they do not require inputs besides the retinal-image data. Their utility, however, can be compromised by the following.

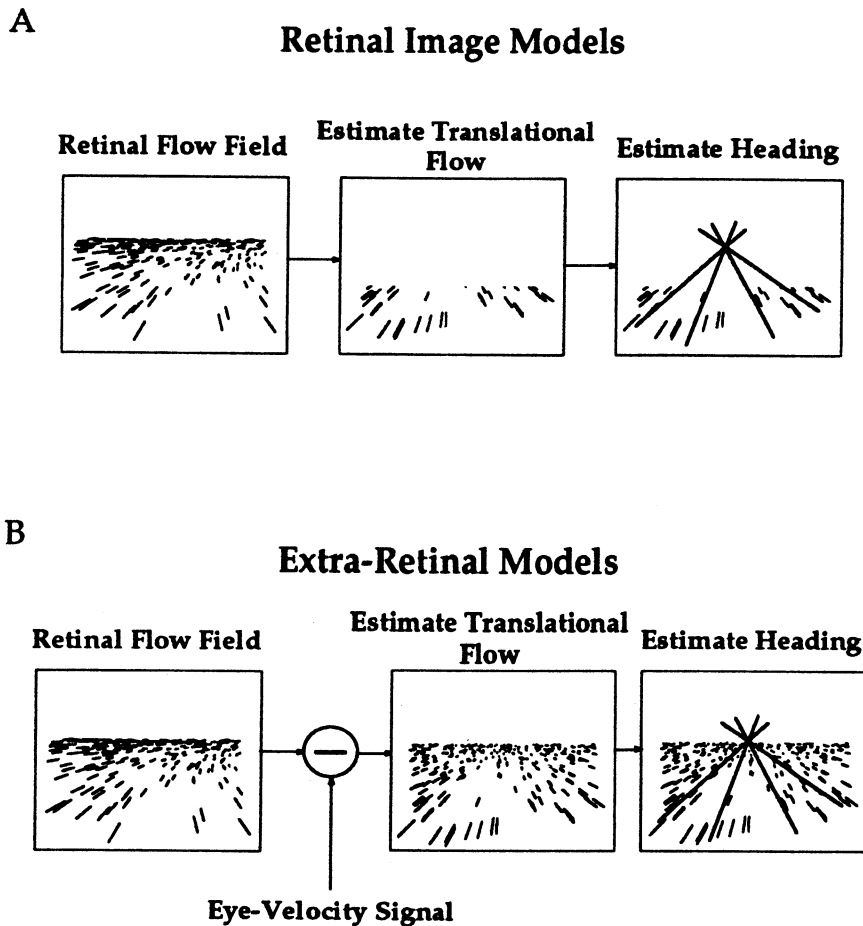


FIGURE 2. Schematics of retinal-image and extra-retinal models. In each case, the input is a translational plus rotational flow field and the output is an estimate of the observer's direction of translation. (A) Retinal-image models. The left panel represents the input to the model. Among retinal-image models, several different algorithms are used to partial out rotational flow and estimate the translational flow components, as schematized by the middle panel. The right panel represents the estimation of heading from the estimated translational flow components. (B) Extra-retinal models. Again the left panel represents the input, the middle panel the estimate of the translational flow component, and the right panel the estimate of the heading. Extra-retinal, eye-velocity signals are used to estimate and subtract the rotational flow component.

- (1) Because of the dependence on depth variation, they do not estimate heading accurately in the presence of rotations when the scene consists of a frontoparallel plane and, to a lesser extent, when the scene contains little depth variation (e.g. Rieger & Lawton, 1985; Koenderink & van Doorn, 1987; Hildreth, 1992).
- (2) Heading estimation is inaccurate when the rotational flow components are large relative to the translational components (Koenderink & van Doorn, 1987; Hildreth, 1992).
- (3) Estimation is inaccurate with small fields of view unless the scene contains sufficient depth variation (Koenderink & van Doorn, 1987).
- (4) Retinal-image-based solutions, by themselves, provide no information about the observer's heading in head- or body-centered coordinates.*

*In fact, there is sufficient retinal image information in *binocular* viewing to determine heading with respect to the head. This point will be taken up in a future paper.

†The role of extra-retinal signals is somewhat different in our discussion of the computation of heading than in the traditional literature on eye movements and spatial vision. An example of the traditional view is provided by Howard's (1982) review; he discusses how extra-retinal signals regarding eye movements can be used to infer object motion in exo-centric coordinates. In our discussion, we consider static visual scenes only, so extra-retinal signals provide information about rotation of the observer's eye and this information is used to distinguish between translational and rotational flow in the optic flow field. The two roles of extra-retinal signals are not mutually exclusive. Future research should consider how heading is perceived in non-static as well as static visual scenes.

In contrast, extra-retinal models estimate the rotational components, R_x , R_y and R_z , directly by means of an extra-retinal eye-velocity signal; the signal could be provided by proprioceptive feedback from the extra-ocular muscles or by efferent signals to those muscles (von Holst, 1954; Matin, 1982).† Assuming that r_x and r_y can be determined accurately from those signals, it is a relatively simple matter to subtract the indicated rotational flow, compute t_x and t_y , and then estimate the heading [see equation (2)].

These models offer at least three advantages to the retinal-image solutions.

- (1) The accuracy of the extra-retinal signals, and thus the accuracy of estimating the rotational components, should not depend on the scene geometry or the field of view.
- (2) Frequently, solutions to the flow equations yield two interpretations of observer motion with respect to a plane (e.g. Longuet-Higgins, 1984). Use of an extra-retinal signal to estimate the observer's rotation provides an easy way around this ambiguity.
- (3) The extra-retinal signal provides information about the position and velocity of the eye with respect to the head thereby allowing a straightforward determination of heading in head-centered coordinates.

There is one major disadvantage to the extra-retinal models. The extra-retinal signals about eye velocity are sometimes erroneous and such errors could yield imprecise or even biased judgments of heading. Under reduced cue conditions, for example, the visual system appears to underestimate the speed of eye rotation during smooth pursuit (Filehne, 1922; Mack & Herman, 1973, 1978). The estimation error (as a percentage of the eye movement velocity) decreases when pursuit velocities are $>1\text{--}2$ deg/sec (Hansen, 1979) or when a large, textured background is present (Mack & Herman, 1978; Wertheim, 1990).

In sum, there are two approaches to recovering heading in the presence of rotations. Given that the retinal-image approach is limited primarily by the visual input and the extra-retinal approach is limited by the accuracy of non-retinal eye-velocity (or even head-velocity) signals, a robust system would estimate heading both ways. When the two estimates differ, the system would choose the one that generally provides more accurate estimates for the current viewing conditions.

Recently, a handful of psychophysicists have investigated whether we rely on one or both of these means of solution. Most notably, Warren and Hannon (1988, 1990) developed a technique for distinguishing the use of retinal-image and extra-retinal algorithms. In one condition of their experiments—the *real eye movement* condition—observers judged their heading while making an eye movement to track a point on a simulated ground plane. In the second *simulated eye movement* condition, observers fixated a stationary point and the flow field deformed to simulate the effects of a tracking eye movement. Figure 1(A, B) depicts the flow fields on the display screen in the real and simulated eye movement

conditions, respectively. To the degree that observers tracked the moving point accurately, the two conditions produced identical flow fields on the retina.* Consequently, retinal-image models predict no difference in performance between the two conditions. Because the eyes did not move in the simulated condition, extra-retinal models predict much poorer performance in that condition.

Warren and Hannon (1988, 1990) reported that observers could reliably distinguish headings differing by about 1.5 deg in both the real and simulated eye movement conditions. They concluded, therefore, that human observers do not require extra-retinal information to judge heading in the presence of eye movements. Their data provide clear support for the biological plausibility of the retinal-image computational models. However, they used very slow pursuit velocities of 0.2–0.7 deg/sec, velocities at which extra-retinal signals are less accurate indicants of rotation speed (Mack & Herman, 1973, 1978). While these may be typical eye movement velocities for someone walking and tracking a distant object, there are many situations in which people make much faster eye and head movements while moving through the world. These situations include tracking a moving object or a near stationary object while walking, running, or driving. For example, a person of average height walking at a normal pace of 1.6 m/sec must make eye movements of 40–50 deg/sec to track texture as it passes under his feet and about 20 deg/sec to track texture 2–3 m in front. Human observers can track objects smoothly at speeds up to 30 deg/sec (Lisberger, Morris & Tyche, 1987), but there is some evidence that they cannot locate heading accurately at such high rotation rates in a simulated eye movement condition (Rieger & Toet, 1985). To explore the ability to solve the rotation problem in greater detail, we measured human observers' heading judgments across a range of eye movement velocities (see also Royden, Banks & Crowell, 1992).

GENERAL METHODS

Four observers, all with corrected visual acuities of 20/20 or better, participated in these experiments. Two of them—MSB and CSR—had considerable experience as psychophysical observers and were aware of the experimental hypotheses. The other two—IVL and TRC—had some experience and were unaware of the hypotheses.

The observers viewed displays of randomly-placed dots whose motions simulated translation or translation plus rotation toward a variety of surfaces. Observers were instructed to fixate a cross that either remained stationary in the center of the display (simulated eye movement condition) or moved horizontally across the display on the horizontal midline (real eye movement condition). Observers viewed the displays monocularly at a distance of 30 cm. The room was completely dark

*Eye rotations cause a displacement of the entrance pupil, so eye movements actually produce translational flow components (Bingham, 1993). However, the translational components are vanishingly small at the simulated distances of our experiments, so one can ignore them.

except for the display. The dots were single pixels subtending 2.9 min arc and did not change size during a motion sequence. In most of the experiments, the dot displays subtended 30×30 deg; this region was defined by a software clipping window. Dots that moved beyond the clipping window disappeared from the display and were not replaced. The stimuli were generated on a TAAC-1 applications accelerator with a Sun 3/160 host and presented on a 16-in. Sony Trinitron monitor. Stimulus frames were drawn to the screen at 33 Hz, half the frame rate of the monitor. The displays yielded a sensation of self-motion.

Before each trial, the first frame of the forthcoming motion sequence appeared until the observer initiated the sequence with a button press. In the real eye movement condition, the fixation point started moving 200 msec before the motion sequence so the observers could establish smooth pursuit before the dots moved. Four types of surfaces, schematized in Figs 3, 6, 8, and 12, were presented in the different experiments. The simulated surfaces generally extended well beyond the limits of the clipping window.

We used two psychophysical procedures to estimate observers' abilities to judge heading. In the two-alternative, forced-choice paradigm (2AFC), a vertical line appeared at the end of the motion sequence and observers indicated whether their heading was to the left or right of the line. No feedback was provided. In order to estimate threshold, the angle between the heading and the target line was varied according to a two-down, one-up, staircase procedure. The staircase was terminated after 12 reversals and an estimate of the 71% correct discrimination threshold was derived from the average of the angles between the target line and true heading for the last 10 staircase reversals. Separate staircases were run for each condition, but they were randomly interleaved in the experimental sessions.

In the seven-alternative, forced-choice paradigm (7AFC), seven equally-spaced lines appeared at the end of the motion sequence and observers indicated the line closest to their perceived heading. No feedback was provided. Unless otherwise stated, the lines were 4 deg apart and therefore spanned 24 deg. Each condition was presented 20 times randomly intermixed with the other conditions. The average of the 20 responses provided the estimate of perceived heading.

EXPERIMENT 1

We first sought to determine whether heading judgments remained accurate at moderate rotation rates with real and simulated eye movements.

Method

The random-dot display simulated translation toward two transparent frontoparallel planes. The duration of each motion sequence was 800 msec, more than double the time needed for most accurate performance in the case of pure translation (Crowell, Royden, Sekuler, Swenson & Banks, 1990). At the start of a trial, the

simulated distances to the planes were 200 and 800 cm. The translational component of the observer motion was 100 cm/sec and the direction varied with uniform probability between extreme values of 5 deg to the left and right of screen center. Two sets of five interleaved staircases were presented for rotation rates of 0, ± 1.25 , 2.5, 5, and 10 deg/sec (a rate of ± 0.62 deg/sec was also

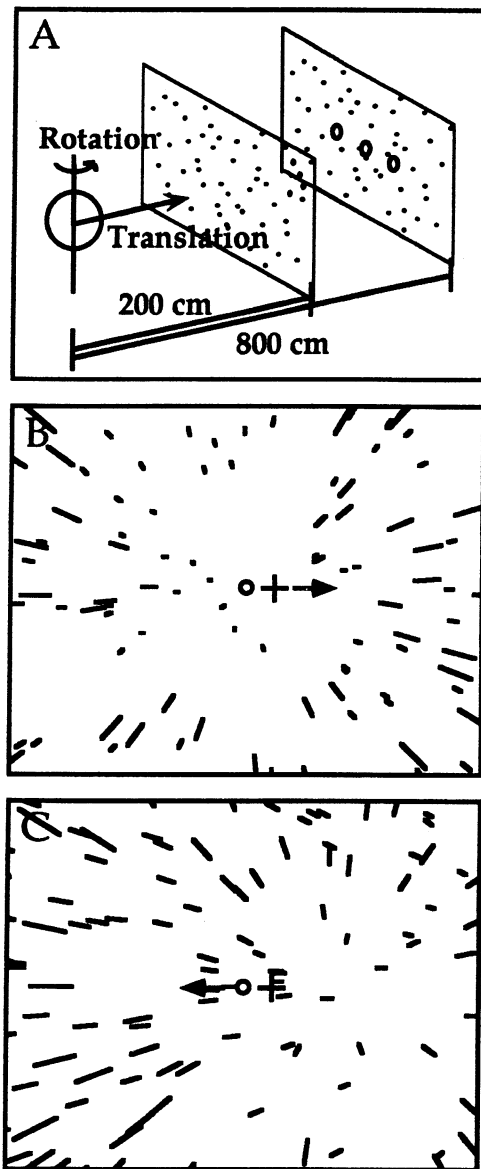


FIGURE 3. Stimuli and flow fields for Expts 1 and 2. (A) Schematic of the scene and observer motion parameters. The simulated scene consisted of two frontoparallel planes at initial distances of 200 and 800 cm. Observer translation at 50 or 100 cm/sec was in one of three directions indicated by the circles. Observer rotation was about a vertical axis. (B) Flow field at the display screen in the real eye movement condition. The field shown here is for translation toward the center of the screen (indicated by the small circle) while the observer makes a 5 deg/sec eye movement to the right (to track the cross). (C) Flow field at the display screen in the simulated eye movement condition. The translation is initially toward the center of the screen (indicated by the circle) and the simulated rotation corresponds to a 5 deg/sec eye movement to the right. To the degree that observers tracked the fixation point accurately, the flow fields at the retina would be identical in the conditions depicted by (B) and (C).

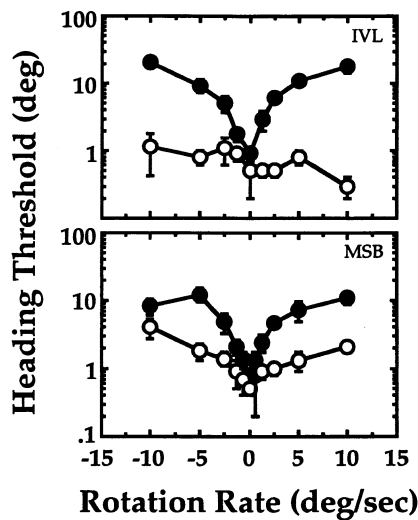


FIGURE 4. Results for Expt 1. Heading discrimination thresholds are plotted as a function of rotation rate for observers IVL and MSB. The stimulus represented approach toward two frontoparallel planes. The open symbols represent thresholds in the real eye movement condition and solid symbols thresholds in the simulated condition. Error bars represent \pm SD. Thresholds were obtained using a 2AFC discrimination procedure without feedback.

presented to observer MSB); Fig. 3 illustrates the flow fields associated with the real and simulated eye movement conditions in this experiment. The 2AFC procedure was used to estimate heading discrimination thresholds.

Results

The results are displayed in Fig. 4. Discrimination thresholds were reasonably small in the real eye movement condition. For observer IVL, thresholds were nearly constant at 0.5–0.7 deg for rotation rates of 0–10 deg/sec; for MSB, thresholds rose from 0.7 to 3.4 deg for the same range of rotations. Thus, when the eyes were moved, heading discrimination was reasonably precise, more so for observer IVL, even at the highest rotation rates tested.

In the simulated eye movement condition, on the other hand, discrimination thresholds rose substantially with increasing rotation rates. For example, thresholds at 10 deg/sec rose to 9.6 and 19.0 deg for MSB and IVL, respectively. Judgments were reasonably accurate, however, at rotation rates of 0.62 and 1.25 deg/sec, values similar to those used by Warren and Hannon (1990). For example, the average threshold at 0.62 deg/sec was 1.4 deg, the same as the value reported by Warren and Hannon for rates of 0.2–1.2 deg/sec.

Clearly, heading discriminations at rotation rates $>$ 1.25 deg/sec were less accurate in the simulated than in the real eye movement condition. This observation is consistent with the predictions of the extra-retinal model. Interestingly, at the higher rates of simulated eye rotation, both observers experienced apparent self-motion along a curvilinear path rather than along a linear path while making an eye movement; in addition, one plane often appeared to move with respect to the

other. The fact that observers' heading judgments were reasonably accurate at slow rotation rates in the simulated condition leaves open the possibility that the visual system can solve the rotation problem without extra-retinal signals under some conditions.

EXPERIMENT 2

In the simulated eye movement condition of Expt 1, observers perceived movement on a curved path instead of the simulated linear movement with a simulated eye rotation. As a consequence, the perceived direction of motion was frequently far from the target line that appeared at the end of the trial. Because of this, the discrimination thresholds probably did not adequately reflect the magnitude of the errors in the observers' heading estimates. In Expt 2, we adopted the 7AFC procedure in order to allow the observers to indicate perceived heading more directly.

Method

As in Expt 1, the random-dot display simulated translation toward two transparent frontoparallel planes at initial distances of 200 and 800 cm. The speed of translation was 50 cm/sec and the directions were straight ahead, 4 deg to the left, or 4 deg to the right. The fixation point was positioned on the horizontal midline of the display. Real and simulated eye rotations were presented with rates of 0, \pm 2.5, or 5 deg/sec. The duration of each motion sequence was lengthened to 1250 msec. At the end of a stimulus presentation, seven vertical lines appeared, spaced 4 deg apart, and the observers indicated the one closest to the perceived heading. The perceived heading was estimated from the average of the observer's responses to 20 stimulus presentations.

Results

The results are shown in Fig. 5. Both observers responded quite differently in the two conditions. The average errors for the real eye movement condition were 0.8 and 1.8 deg for the 2.5 and 5 deg/sec rotation rates, respectively. The corresponding average errors in the simulated condition were 7.5 and 15.5 deg. The value at 5 deg/sec in the simulated condition may actually under-represent the magnitude of the perceptual errors because observer responses were constrained to the span of the target lines (\pm 12 deg from screen center). It does not appear, however, that these data were highly constrained by such a ceiling effect because neither observer chose the leftmost or rightmost target line more than 75% of the time for any condition.

These results support the earlier conclusion that the visual system uses extra-retinal information about eye movements to aid the estimation of heading in the presence of rotations. If the extra-retinal information signals stationary eyes, as in the simulated eye movement condition, then observers incorrectly perceive self-motion along a curved path.

EXPERIMENT 3

As mentioned earlier, retinal-image models capitalize on depth variation in the scene to recover the translational component of the observer's motion; thus, they have difficulty estimating heading in the presence of rotations for translation perpendicular to a single plane (e.g. Koenderink & van Doorn, 1987). We wondered if the availability of an extra-retinal eye-velocity signal would allow accurate estimates in this situation.

Method

The observers, procedure, and stimuli in Expt 3 were the same as in Expt 2 with one exception: the simulated scene consisted of a single frontal plane with an initial distance of 200 cm. The scene and resulting flow fields are presented schematically in Fig. 6.

Results

The results, displayed in Fig. 7, were very similar to those of Expts 1 and 2. Heading judgments remained accurate in the real eye movement condition and inaccurate in the simulated condition. In the real eye movement condition, the average errors were 1.5 and 1.9 deg for rotation rates of 2.5 and 5 deg/sec, respectively; in the simulated condition, the corresponding errors were 9.5 and 17.1 deg. One observer's performance in the real and simulated conditions was similar for a rotation rate of 1 deg/sec. The displays in this

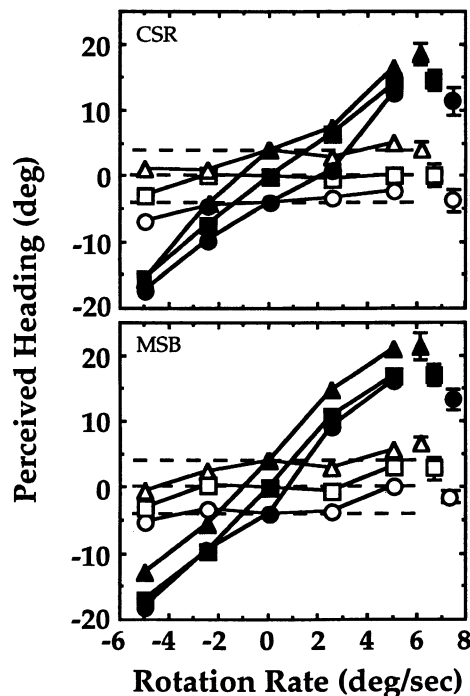


FIGURE 5. Results for Expt 2. Perceived heading is plotted as a function of rotation rate for observers CSR and MSB. The stimulus represented approach toward two frontoparallel planes. Open symbols represent responses in the real eye movement condition and solid symbols responses in the simulated condition. Dotted lines indicate the actual headings of -4 , 0 , and 4 deg, and the circles, squares, and triangles represent the responses for those headings, respectively. Error bars on the right show twice the average SD for each heading.

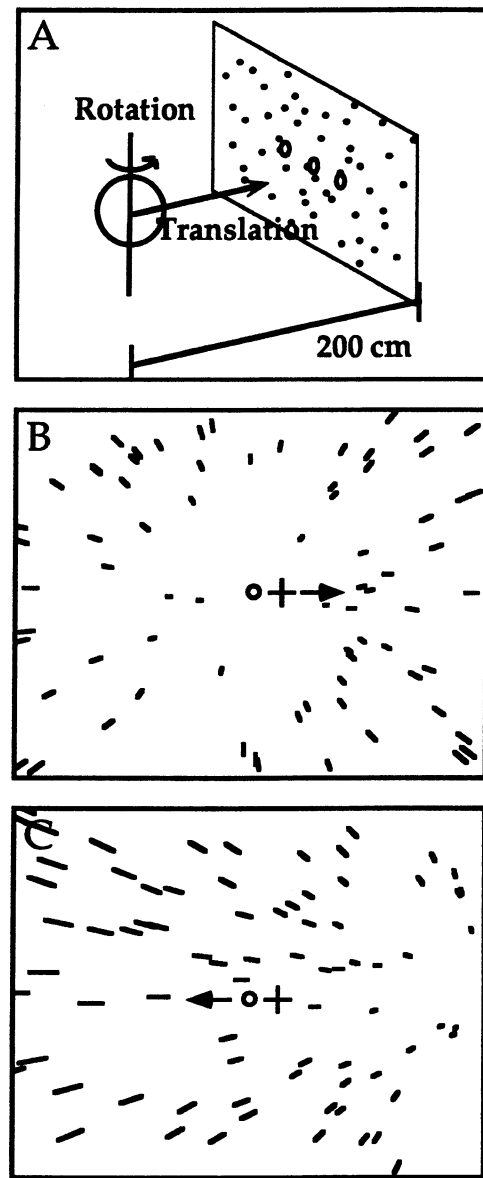


FIGURE 6. Stimuli and flow fields for Expt 3. (A) Schematic of the scene and observer motion parameters. The simulated scene consisted of one frontoparallel plane at an initial distance of 200 cm. Observer translation at 50 cm/sec was in one of three directions indicated by the circles. Observer rotation was about a vertical axis. (B) Flow field at the display screen in the real eye movement condition. The field shown here is for translation toward the center of the screen (indicated by the small circle) while the observer makes a 5 deg/sec eye movement to the right (to track the cross). (C) Flow field at the display screen in the simulated eye movement condition. The translation is initially toward the center of the screen and the simulated rotation corresponds to a 5 deg/sec eye movement to the right. The flow fields at the retina would be identical in the conditions depicted by (B) and (C).

experiment produced apparent foci of expansion in the retinal image that did not correspond with the heading (see Fig. 6). The positions of these foci at trial end are indicated by the diagonal dashed lines in Fig. 7. The close correspondence between these lines and the simulated eye movement data show that observers chose headings that coincided with the apparent foci in the simulated condition.

The results of Expt 3 show that, when extra-retinal

signals for eye velocity are available, human observers make accurate heading judgments in a situation in which retinal-image models generally do not.

EXPERIMENT 4

The results of Expts 1–3 suggest that information from the execution of an eye movement is frequently useful in solving the rotation problem. Because this observation seems inconsistent with Warren and Hannon's (1988, 1990) data, we sought to determine the critical differences between our experiments and theirs. Most of Warren and Hannon's displays simulated an observer walking across a ground plane, so we asked whether the rotation problem is handled differently in this situation as compared to the situations simulated in Expts 1–3. To do this, we reproduced the conditions of Warren and Hannon (1988) and the ground plane experiments of Warren and Hannon (1990) in most respects.

Method

The dot motions corresponded to the flow produced when an observer translates across a ground plane at a speed of 190 cm/sec. Simulated eye height was 160 cm. The ground plane was truncated at a distance of 3730 cm. Figure 8 schematizes the surface and flow

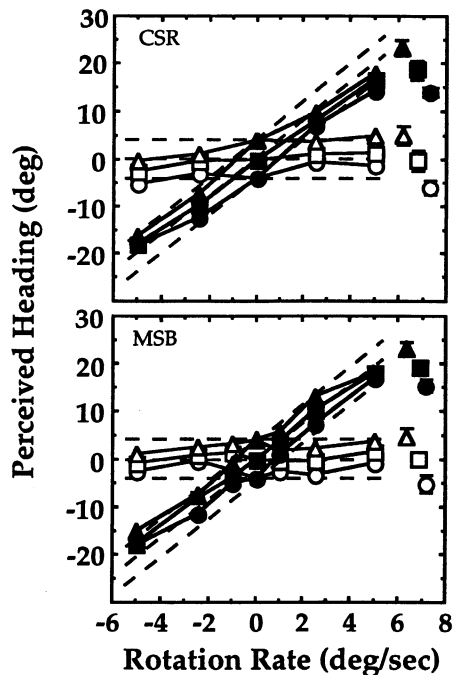


FIGURE 7. Results for Expt 3. Perceived heading is plotted as a function of rotation rate for observers CSR and MSB. The stimulus represented approach toward one frontoparallel plane. Open and solid symbols represent responses in the real and simulated condition, respectively. Horizontal dotted lines indicate the actual headings of -4 , 0 , and 4 deg, and the circles, squares, and triangles represent the responses for those headings, respectively. Error bars on the right show twice the average SD for each heading. Diagonal dotted lines show the positions of the apparent focus of expansion (i.e. the position of zero velocity in the flow field) at the end of the motion sequence for different conditions.

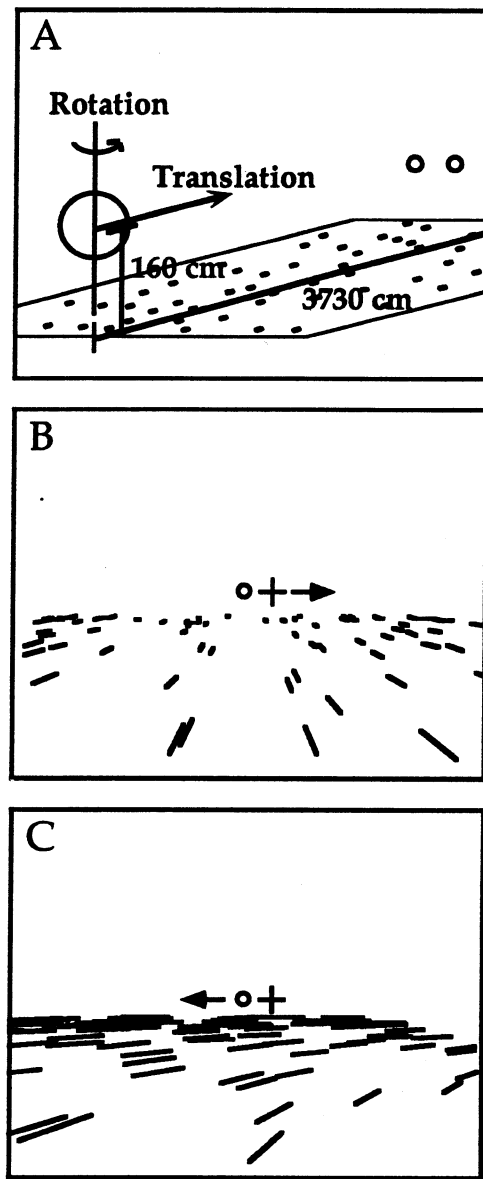


FIGURE 8. Stimuli and flow fields for Expt 4. (A) Schematic of the scene and observer motion parameters. The simulated scene consisted of a ground plan extending to 3730 cm. Simulated eye height was 160 cm. Observer translation was parallel to the ground plane at 190 cm/sec and was in one of three directions indicated by the circles. The observer fixated a point 2.5 deg above the truncated horizon. The point moved horizontally so rotation was about a vertical axis. Dots were distributed randomly on the plane with an average density of 0.6 dots per m^2 ; roughly 220 dots were visible at the beginning of a trial. (B) Flow field at the display screen in the real eye movement condition. The field shown here is for translation toward the center of the screen (indicated by the small circle) while the observer makes a 5 deg/sec eye movement to the right (to track the cross). (C) Flow field at the display screen in the simulated eye movement condition. The translation is initially toward the center of the screen and the simulated rotation corresponds to a 5 deg/sec eye movement to the right. The flow fields at the retina would be identical in the conditions represented by (B) and (C).

fields. Trial duration was 1250 msec. The direction of translation was straight ahead, 4 deg to the left or 4 deg to the right of screen center. Experiment 4 differed from Warren and Hannon's experiments in the following ways. First, we used constant, faster rotation rates (real

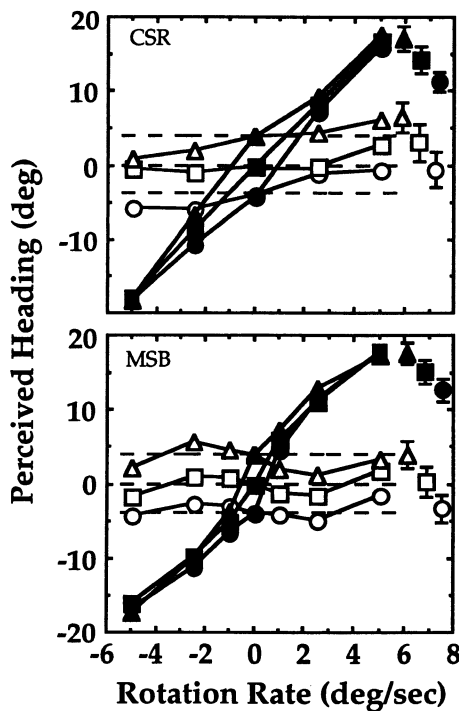


FIGURE 9. Results for Expt 4. Perceived heading is plotted as a function of rotation rate for observers CSR and MSB. The stimulus represented translation parallel to a ground plane. Open and solid symbols represent responses in the real and simulated condition, respectively. Horizontal dotted lines indicate the actual headings of -4 , 0 , and 4 deg and the circles, squares, and triangles represent the responses for those headings, respectively. Error bars on the right show twice the average SD for each heading.

or simulated) and a vertical axis of rotation; in their experiments, observers tracked a point on the ground plane (resulting in an increasing rotation rate over the course of the trial) and the point's location was chosen randomly from trial to trial (resulting in a variety of rotation axes). Second, the fixation point was positioned 2.5 deg above the horizon and moved independently of the ground plane; it was attached to the ground plane in their experiments. Third, we used the 7AFC procedure to estimate perceived heading in the various conditions; they used the 2AFC procedure. We tested one observer at rotation rates of ± 1 deg/sec in which case the target lines appearing at the end of the trial were spaced 2 rather than 4 deg apart.

Results

Figure 9 shows the results. Once again the observers reported different headings depending on whether the rotational flow component was created by a real or simulated eye movement. The average errors in the real eye movement condition were 1.5 and 1.9 deg for the 2.5 and 5.0 deg/sec rotation rates, respectively. The corresponding errors in the simulated condition were 9.8 and

17.3 deg. In addition, responses at slow rotation rates of 1 deg/sec were significantly more accurate in the real eye movement condition; this specific conclusion is limited, however, by the fact that only one observer was presented rotation rates less than 2.5 deg/sec. These observations imply that the type of simulated scene was not the critical difference between our experiments and Warren and Hannon's (1988, 1990).

EXPERIMENT 5

Because the results of Expts 1–4 conflicted with Warren and Hannon's (1988, 1990), we performed an exact replication of their ground plane experiment.

Method

The dot motions corresponded to the flow produced when an observer translates across a ground plane at a speed of 190 cm/sec with an eye height of 160 cm. The ground plane was truncated at a distance of 3730 cm and the fixation target was attached to the ground plane; as in Warren and Hannon's experiments, the position of the point varied randomly from trial to trial, so the average rotation rate varied from 0.2 – 0.7 deg/sec.* Again, real and simulated eye movement conditions were presented. We used the 2AFC procedure to estimate heading discrimination thresholds.

Results

As reported by Warren and Hannon (1988, 1990), discrimination thresholds were 1 – 2 deg in both the real and simulated eye movement conditions. Therefore, observers were able to judge heading reasonably accurately under these conditions whether they moved their eye or not.

EXPERIMENT 6

In Expts 1–4 we found that observers generally could not estimate heading accurately in the presence of rotations > 1 deg/sec when the rotation was created by a simulated rather than an actual eye movement. However, the observers in Expt 5 judged heading rather accurately in the simulated condition in an exact replication of Warren and Hannon's ground plane experiments. More importantly, van den Berg (1993) has reported that heading estimation is reasonably accurate for simulated rotation rates of nearly 5 deg/sec when the display mimics observer translation parallel to a ground plane while fixating a point on the ground. van den Berg's observation leads to the question, what was the critical difference between our displays in Expts 1–4 and his? There are two possibilities.

- (1) Attachment of the fixation target to the simulated scene. In Expts 1–4, the fixation target was not attached to the otherwise rigid scene; in van den Berg's (and Warren and Hannon's experiments and Expt 5), the target was a point on the ground plane. Perhaps

*Observers in Warren and Hannon's experiments were asked to make left–right judgments of their heading with respect to a vertical target line. Thus, only the horizontal component of the rotation is relevant to the judgment. For this reason, the range of relevant rotation rates was smaller than the reported range.

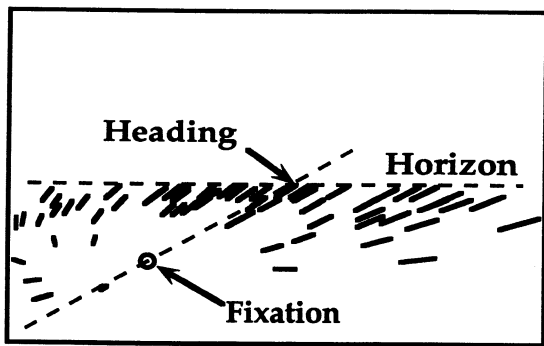


FIGURE 10. Depiction of the horizon cue. The flow field represents the pattern of retinal image motions corresponding to translation parallel to a ground plane while making an eye movement to track a point on the ground (indicated by the small circle). The direction of translation is indicated by the vertical line at the horizon. For translation parallel to a ground plane, the heading is always located at the horizon and the fixation point always moves away from the heading. Thus, the heading corresponds to the intersection of the horizon with a line of common flow directions through the fixation point regardless of the axis and rate of eye rotation.

heading is estimated more readily in the presence of rotations when the observer is tracking a point attached to the rigid scene.

- (2) Use of a horizon cue. Translation across a ground plane (i.e. translation parallel to a plane) provides a simple cue that is not present in most other situations. As shown in Fig 10, heading corresponds with the intersection of the horizon and a line of common flow directions (van den Berg, 1992). Perhaps observers capitalized on this horizon cue in van den Berg's (1993) displays.

We tested these hypotheses in Expts 6 and 7. The first examined the use of the horizon cue.

Method

The displays and procedure were identical to those of Expt 5 with the following exceptions. (1) The fixation point, which was a point on the ground plane, was positioned along a 45-deg diagonal through the screen center. The axis of rotation, therefore, was diagonal rather than variable as it was in Expt 5. (2) The simulated headings were also along the diagonal in three possible locations: screen center, 4 deg down and to the left, or 4 deg up and to the right. In other words, the simulated translations were not constrained to be parallel to the ground plane (i.e. the XZ plane) as they were in Expt 5 and in Warren and Hannon's and van den Berg's (1993) ground plane experiments. Adding a varying translational component within a plane perpendicular to the axis of rotation eliminates the horizon cue (van den Berg, 1992). (3) The seven target lines appeared along the same diagonal. Preliminary observations suggested that heading judgments were less accurate in the simulated eye movement condition, so the target lines were spaced 6 deg apart in the simulated eye movement condition and 4 deg apart in the real condition. (4) The display was slightly larger at 42×32 deg.

Because the simulated translational speed and eye height were fixed at 190 cm/sec and 160 cm, respectively (see Expts 4 and 5), the rotation rate was determined by the locations of the fixation point and simulated heading. In the results reported below, the reported rotation rate is the average over the course of the trial.

Results

The results are displayed in Fig. 11. The observer again reported different headings depending on whether an executed or simulated tracking eye movement caused the rotational flow. The average errors in the real eye movement condition were 1.2, 1.1, and 1.7 deg for rotation rates of 1, 2.5, and 5 deg/sec, respectively. The corresponding errors in the simulated condition were 2.0, 10.0, and 16.0 deg. The errors in the simulated condition were substantially higher at rotation rates greater than 1 deg/sec. Therefore, when the horizon cue was unavailable, observers did not recover heading accurately in the presence of fast rotational flow not accompanied by an eye movement. Thus, hypothesis 2 above cannot be rejected; observers may have judged heading accurately in the simulated eye movement conditions of Expt 5 and Warren and Hannon's (1988, 1990) and van den Berg's (1993) ground plane experiments by using the horizon cue.

EXPERIMENT 7

We next examined the first of the above-mentioned hypotheses: is attachment of the fixation point to the simulated rigid scene an important variable in determin-

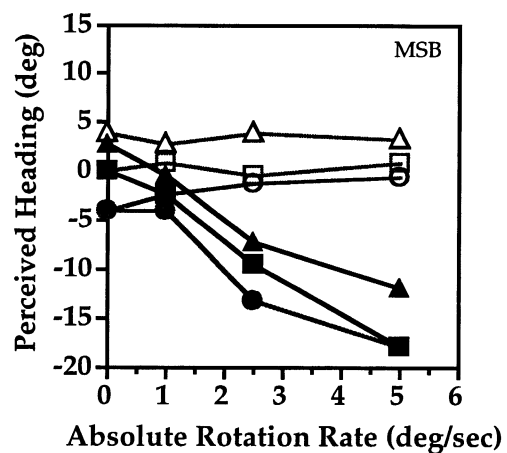


FIGURE 11. Results for Expt 6. Perceived heading is plotted as a function of rotation rate for observer MSB. The stimulus represented translation relative to a ground plane; the horizon cue was eliminated because the translation involved a component perpendicular to the ground plane. The rotations were negative (i.e. to the left and down), but have been plotted as positive numbers for simplicity. The rotation rates shown are magnitudes along the diagonal described in the text. Open and solid symbols represent responses in the real and simulated condition, respectively. Horizontal dotted lines indicate the actual headings of -4 , 0 , and 4 deg and the circles, squares, and triangles represent the responses for those headings, respectively.

ing whether observers can solve for the heading in the presence of simulated eye rotations? We eliminated the horizon cue by simulating translation through a three-dimensional cloud of randomly-positioned dots with a fixation point that was a member of the cloud (thereby attaching fixation to the simulated rigid object).

Method

The motion of the dots in the display simulated translation at 250 cm/sec through a cloud of dots at distances of 0–3730 cm. Roughly 615 dots were visible at the beginning of the trial. The simulated scene and flow fields are schematized in Fig. 12. The fixation point was a member of the rigid cloud and was positioned 5 deg to the left or right of the heading at the beginning of the motion sequence. Placing it at initial distances of 10⁷, 1560, 810, and 560 cm produced average rotation rates of 0, 1, 2.5, and 5 deg/sec, respectively. The direction of translation was straight ahead, 4 deg to the left, or 4 deg to the right of screen center. Stimulus duration was 1250 msec and the displayed flow field subtended 30 × 30 deg. All other display parameters were identical to those of the previous experiments. The 7AFC procedure was used to estimate perceived heading for the various experimental conditions.

Results

The results are displayed in Fig. 13. Once again, observers judged heading accurately when they actually moved their eye at rates > 1 deg/sec and inaccurately when such eye movements were simulated. Both observers performed similarly in the real and simulated conditions for rotations of 1 deg/sec or less. There are two reasonable interpretations of the similarity of performance at slow rotation rates: (1) observers may have estimated heading accurately (more so for MSB than for TRC) from retinal image information alone at slow rates (Warren & Hannon, 1988, 1990) and (2) observers may not have estimated heading accurately at slow simulated rates but the target lines were spaced too far apart to detect the perceptual errors. We cannot distinguish between these possibilities currently.

The fact that observers judged heading with reasonable accuracy in the simulated condition for slow, but not fast, rotation rates is consistent with the idea that speed of rotation is critical.* The horizon cue was unavailable in this experiment, so the observation of reasonably accurate judgments at simulated rotation rates of 1 deg/sec implies that this cue is not critical at slow rotation rates. The fixation point was attached to the simulated object, so for rotation rates above 1 deg/sec, the data of Fig. 13 rule out hypothesis 1 above which states that attachment of the fixation point to the scene is the critical variable. For rotation rates of 1 deg/sec or less, the data do not differ substantially from those of Expt 4 in which the fixation point was not

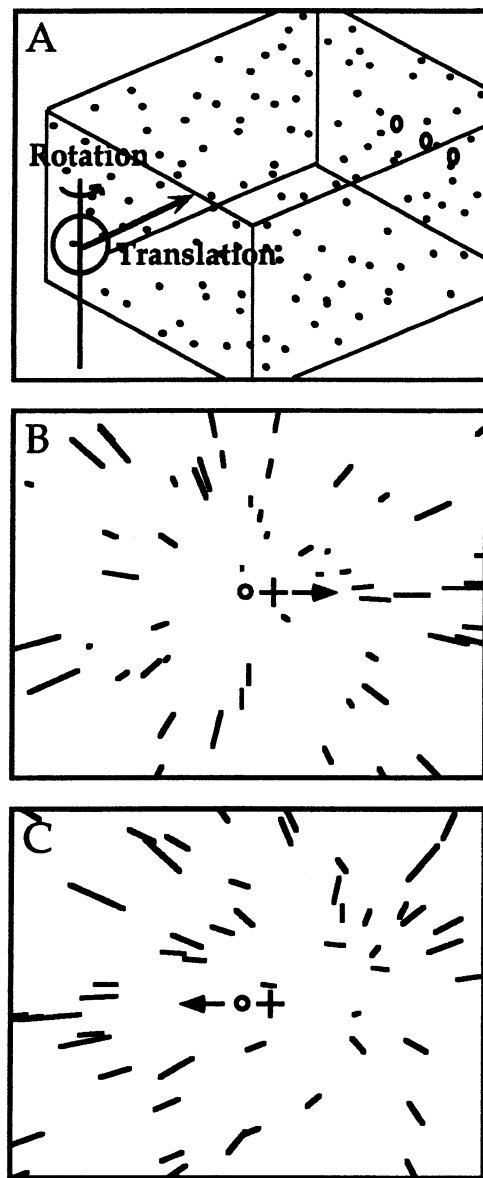


FIGURE 12. Stimuli and flow fields for Expt 7. (A) Schematic of the scene and observer motion parameters. The simulated scene was a rigid, three-dimensional cloud of dots; initial distances of the dots were 0–3730 cm. Roughly 615 dots were visible at the beginning of the motion sequence. Observer translation was at 250 cm/sec and was in one of three directions indicated by the circles. The fixation point was a member of the three-dimensional cloud and was positioned 5 deg to the left or right of the heading at the beginning of the motion sequence. Placing it at initial distances of 10⁷, 1560, 810, and 560 cm produced average rotation rates of 0, 1, 2.5, and 5 deg/sec, respectively. (B) Flow field at the display screen in the real eye movement condition. The field shown here is for translation toward the center of the screen (indicated by the small circle) while the observer makes a 5 deg/sec eye movement to the right (to track the cross). (C) Flow field at the display screen in the simulated eye movement condition. The translation is initially toward the center of the screen and the simulated rotation corresponds to a 5 deg/sec eye movement to the right. The flow fields at the retina would be identical in the conditions depicted by (B) and (C).

attached. Thus, at slow rotation rates, the attachment of the fixation point again does not appear to be a critical variable. The results from Expts 6 and 7 imply that speed of rotation and the availability of the horizon (at high rotation rates) are the critical factors in the differences

*It seems likely that the ratio of translation and rotation speeds would be the critical factor (Koenderink & van Doorn, 1987), but these experiments are not designed to test this possibility directly.

we observed in Expts 1–4 as compared to Expt 5 and Warren and Hannon's (1988, 1990) and van den Berg's (1993) experiments.

EXPERIMENT 8

There was an additional cue in Warren and Hannon's (1988, 1990) and van den Berg's (1992, 1993) experiments and in some of the above experiments (Expts 1, 5, and 6) that may have aided performance in the real relative to the simulated eye movement condition. In the real eye movement condition, the fixation point moved relative to the boundaries of the displayed flow field and in the simulated condition, it did not. Because motion of the fixation target with respect to the surroundings normally accompanies eye movements (Gibson, 1966), observers could have used the motion of the fixation point relative to the visible edge of the display in the real movement condition to estimate the velocity of the rotational flow component. Experiment 8 examined the possibility that observers capitalized on this relative motion cue.

Method

The stimuli and procedure were identical to those of Expt 7 with two exceptions.

(1) The software clipping window was either stationary or moving thereby producing four experimental conditions: real eye movement with stationary window

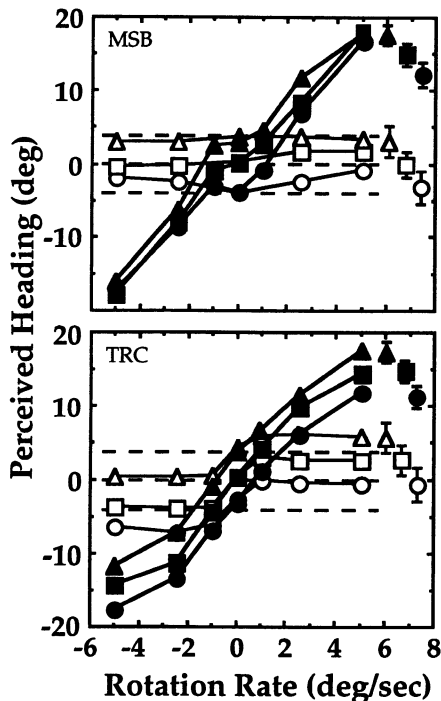


FIGURE 13. Results for Expt 7. Perceived heading is plotted as a function of rotation rate for observers MSB and TRC. The stimulus represented translation through a three-dimensional cloud. Open and solid symbols represent responses in the real and simulated condition, respectively. Horizontal dotted lines indicate the actual headings of -4 , 0 , and 4 deg, and the circles, squares, and triangles represent the responses for those headings, respectively. Error bars on the right show twice the average SD for each heading.

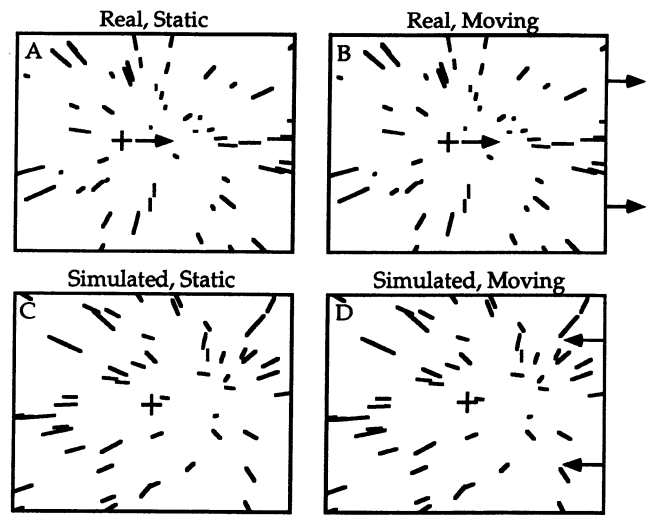


FIGURE 14. Stimuli for Expt 8. The scene consisted of a rigid three-dimensional cloud of dots. The visible portion of the display was determined by a software clipping window. The four conditions are schematized in separate panels that depict the stimulus at the display screen. (A) Real eye movement with stationary clipping window; the observer fixated a moving point during the motion sequence and the visible portion of the screen was stationary with respect to the center of the display screen. (B) Real eye movement with moving clipping window; the observer fixated a moving point and the visible portion of the display screen moved at the same velocity as that point. (C) Simulated eye movement with stationary clipping window; the observer fixated a stationary point and the visible portion of the screen was stationary with respect to the center of the display screen. (D) Simulated eye movement with moving clipping window; the observer fixated a stationary point and the visible portion of the display moved at the same speed as, but opposite direction, from the simulated eye rotation.

(the real eye movement condition of Expt 7), real eye movement with moving window (in which case the window moves at the same velocity as the fixation point), simulated eye movement with stationary window (the simulated eye movement condition of Expt 7), and simulated eye movement with moving window (the window moved at the same speed as, but opposite direction, from the simulated eye movement). These displays are schematized in Fig. 14. Notice that the relationship between the clipping window and eye position is the same with real eye movements and a stationary window as it is with simulated eye movements and a moving window (in both cases, the image of the clipping window moves across the retina). Similarly, the relationship between window and eye position is the same with real eye movements and a moving window as it is with simulated movements and a stationary window (in both cases, the image of the window does not move across the retina).

(2) The displayed flow fields subtended 40×40 deg as compared to the 30×30 deg fields of Expt 7 and most of the other experiments. We used larger displays in order to test the possibility that poor performance in the simulated eye movement conditions was due to the relatively small fields of view available to the observers (Koenderink & van Doorn, 1987; Hildreth, 1992). Unfortunately, we could not present fields larger than 40×40 deg with the equipment available at the time.

Results

The results are displayed in Fig. 15 which plots the absolute values of perceived heading errors averaged across the three simulated headings (-4 , 0 , and 4 deg) and the two rotation rates of the same magnitude but opposite directions (e.g. -5 and 5 deg/sec). Although the effect was larger in observer MSB than in TRC, both observers judged heading more accurately when the rotational flow was created by an executed eye movement than by a simulated movement whether the clipping window moved across the retina or not. This finding suggests that the relative motion of the fixation point and the edge of the displayed flow field does not significantly affect the ability to judge heading in the presence of rotations.

DISCUSSION

Comparison with previous reports

Under most conditions, our observers estimated heading accurately after an executed eye rotation, but not after a simulated rotation. This seems to contradict evidence from several previous studies that human observers can actually estimate heading accurately with simulated eye movements (Rieger & Toet, 1985; van den Berg, 1992, 1993; Warren & Hannon, 1988, 1990). However, with the possible exceptions of van den Berg (1992, 1993), these reports are not inconsistent with our general conclusions. Here we review this work and compare it to our findings.

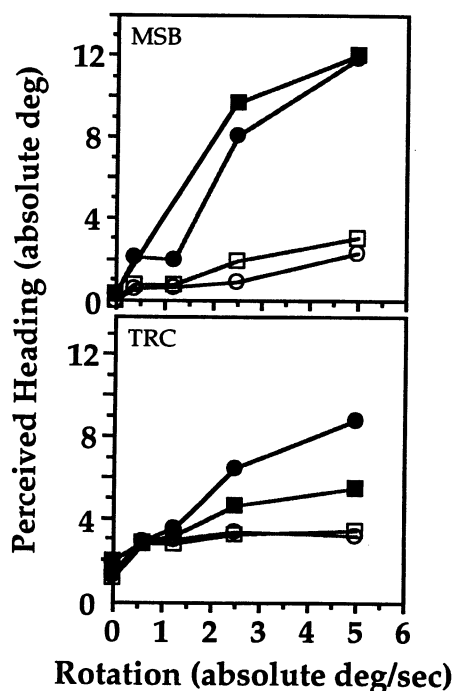


FIGURE 15. Results for Expt 8. Perceived heading is plotted as a function of rotation rate for observers MSB and TRC. The stimulus represented translation through a three-dimensional cloud. Open and solid symbols represent responses in the real and simulated condition, respectively. Squares represent responses with a moving clipping window and circles responses with a stationary window.

Rieger and Toet (1985) performed the first study on the perception of heading with simulated eye rotations. They simulated motion of the observer toward transparent frontoparallel planes; the motion consisted of a translation in a straight line combined with a rotation around a randomly-selected axis (all axes, including the line of sight, having equal probability). The translations were toward one of four points at the vertices of a square centered in the screen. The observers indicated which point corresponded to the perceived direction of translation. Rieger and Toet varied the rotation rate (from 0 to 1.8 deg/sec), the number of planes (one plane at a distance of 3 m or two planes, one at 3 m and one at 12 m), the size of the display (10×10 or 20×20 deg) and the spacing between the four alternative headings (2.5 or 5 deg). When the simulated scene consisted of a single, frontoparallel plane, the percentage of correct responses declined rapidly as the rotation rate increased. If the scene consisted of two planes, however, some observers were able to identify the correct heading with reasonable accuracy up to the highest rotation rate of 1.8 deg/sec. Changing the field of view from 20×20 to 10×10 deg had no effect.

Although differences in the task and axis of rotation make direct comparisons between Rieger and Toet's (1985) and our data difficult, there is considerable agreement. Consistent with our findings (Expts 7 and 8), Rieger and Toet's showed that human observers can identify heading reasonably accurately in the presence of slow simulated rotations. They also observed a deterioration in performance at higher rotation rates (1.0–1.8 deg/sec) even when the scene consisted of two planes and the larger field of view.

Warren and Hannon (1988, 1990) also examined the perception of heading during real and simulated eye movements. In their 1988 paper and in Expt 2 of their 1990 paper, they simulated forward translation of the observer across a ground plane. The observer either tracked a point on the ground (the real eye movement condition) or fixated a stationary point while the dots representing the ground plane moved to simulate tracking a point on the ground (the simulated eye movement condition). The observers indicated whether the simulated translation was to the left or right of a vertical line that appeared on the horizon at the end of each motion sequence. Discrimination thresholds did not differ significantly between the two conditions. In their Expt 3, Warren and Hannon (1990) replaced the ground plane with a random three-dimensional cloud and obtained the same pattern of results.

Warren and Hannon concluded that the visual system can decompose complex flow fields into translational and rotational components on the basis of purely retinal stimulation. Despite this conclusion, their results do not disagree with ours because they used very slow rotations of 0.2–0.7 deg/sec. At rates of 1.0 deg/sec or less, we too observed reasonably accurate heading judgments in the simulated condition. Two additional points should be considered in evaluating Warren and Hannon's (1988, 1990) experiments. First, the rotation required to track

a fixation point on the ground contains components around two axes: horizontal (X) and vertical (Y). The rotational component around the X -axis is usually larger, but only the component around the Y -axis is relevant to a left-right heading discrimination task.* Second, for translation parallel to the ground ($T_y = 0$), a visible horizon provides a simple two-dimensional cue for discriminating headings in the horizontal plane: the heading simply corresponds to the point on the horizon that is moving directly away from the fixation point. Earlier we referred to this as the horizon cue. These criticisms do not apply to Warren and Hannon's (1990) three-dimensional cloud experiment.

Warren and Hannon's (1988, 1990) results, like Rieger and Toet's (1985), indicate that human observers have some ability to recover heading in the presence of rotational flow even when the appropriate extra-retinal signal is missing or contradictory. Both studies used very slow rotations, however, and do not address the question of what happens at rotation rates above 1–2 deg/sec.

van den Berg (1992) performed a series of experiments in which a certain proportion of dots (the "noise" dots) moved in random directions in order to insert noise into the flow field. His Expt 2, which was similar to the ground plane experiments of Warren and Hannon (1988, 1990), revealed that heading discrimination was equally good in real and simulated eye movement conditions when almost all the dots were "signal" dots. However, as the "signal-to-noise" ratio decreased (by decreasing the proportion of "signal" dots), discrimination performance in the simulated condition deteriorated rapidly while performance in the real eye movement condition was unaffected. Experiments 3 and 5 in van den Berg (1992) yielded basically the same pattern of results in the simulated condition even when the horizon cue present in Expt 2 was eliminated. Unfortunately, one cannot determine the rotation rates presented in van den Berg's experiments; he simply states that they were always <3.2 deg/sec. Additionally, the rotations in van den Berg's experiments had significant vertical rotation components (i.e. rotations about the X -axis) that were not relevant to the observers' horizontal discrimination task in Expts 2 and 5. We cannot, therefore, compare van den Berg's and our results directly because we cannot determine the magnitudes of the rotational components nor the proportions of the components that were in the direction relevant to the observers' task.

An experimental observation of van den Berg (1993) conflicts with our general conclusion that the execution of an eye movement matters at all but slow rotation rates. Specifically, he reported that observers could judge heading reasonably accurately for simulated rotations up to roughly 5 deg/sec. However, van den Berg's displays provided a horizon cue and the results of

Expt 6 showed that observers do not estimate heading accurately at high simulated rotation rates once the cue is eliminated. Thus, our general conclusions appear to be valid except for conditions that provide a usable horizon cue.

A few investigators (Warren, Mestre, Blackwell & Morris, 1991; Perrone & Stone, 1994) have studied the perception of heading on a circular path in the presence of simulated rotations. Their results are undoubtedly relevant to an understanding of how the visual system solves the rotation problem, but the tasks in those experiments and ours differ too much to allow meaningful comparison.

Usefulness of the horizon cue

Warren and Hannon (1988, 1990), van den Berg (1992, 1993) and we (Expt 5) found that observers could estimate heading reasonably accurately when the motion sequence depicted translation parallel to a ground plane while fixating a point attached to the plane. Recall that this viewing situation creates the horizon cue: that is, the heading corresponds to the intersection of the horizon with a line of common flow directions through the fixation point regardless of the rate of eye rotation. The results of van den Berg (1993) and Expt 5 showed that observers could estimate heading accurately during simulated eye rotations when the horizon cue was available and the results of Expt 6 showed that their estimates were adversely affected when the cue was eliminated.

How frequently is the horizon cue available in everyday situations? This question cannot be answered without a better description of locomoting observers' viewing habits, but three conditions must exist for the cue to be informative. First, the fixation point must be attached to the ground plane (although it need not be at the same elevation). Second, the horizon must be at approximately the same elevation as the ground beneath the observer and must be visible (and obviously distant). Third, the observer's motion cannot contain a significant vertical component. These conditions would be met for an observer fixating a lamp post while walking on flat and open terrain; they would not be met for numerous other situations such as walking through undulating or cluttered terrain or landing an airplane. Thus, the horizon cue cannot be used in many everyday situations.

Use of retinal and extra-retinal information

Our results show that under many, perhaps most, viewing conditions the execution of an eye movement matters in determining the direction of self-motion when the flow field consists of translational and rotational flow. However, one should not conclude from this that the visual system does not use methods like those of the retinal-image models described earlier. In particular, there is an important difference between the predictions of the retinal-image and extra-retinal models for the real and simulated eye movement conditions. In the real eye movement condition, both models receive input that is commonplace for everyday situations because the presumed extra-retinal eye-velocity signal matches the

*Strictly speaking, addition of a rotational component about the horizontal axis probably decreases the signal-to-noise ratio even for a task involving a horizontal discrimination of possible headings, but the magnitude of the effect is much less than for rotational components about a vertical axis (Rieger & Lawton, 1985).

rotational flow in the retinal image [to the degree that it normally matches rotational flow (Mack & Herman, 1973, 1978)]. Thus, both methods of solving the rotation problem should yield similar estimates of heading in the real eye movement condition. With simulated eye movements, however, the inputs are discordant from the perspectives of the two models. In the simulated condition, the retinal-image model receives the same input as it does in the real movement condition, but the extra-retinal model does not; specifically, from the perspective of the extra-retinal model, the eye-velocity signal provides information that the eye has not moved, yet the retinal image contains rotational flow that normally accompanies an eye movement. A cue conflict thus arises that might well affect the observer's response. For this reason, inaccurate judgments in the simulated condition cannot be taken as evidence that methods like the retinal-image models described above are not used in estimating heading. Of course, it remains possible that retinal-image-based solutions are in fact not used at high rotation rates; the point is that our data are not decisive on this issue.

The central issue in this paper—how the visual system deals with the effects of eye movements while estimating the direction of self-motion—has clear connections to issues with a long scientific history. A classical question in visual perception has been: how do we maintain the percept of a stable world during eye movements? There were two classical theoretical approaches to explaining this phenomenon. The inferential theory held that the visual system compares two sources of information—retinal afferent signals and extra-retinal, eye-velocity signals—to determine whether the external scene remained stationary during an eye movement (e.g. Helmholtz, 1910; von Holst & Mittelstaedt, 1950; Sperry, 1950). The direct theory, on the other hand, claimed that the visual system solves the problem from retinal image information alone by picking up certain invariant features of image motions that normally accompany an eye movement (e.g. Gibson, 1966, 1968).

Neither of these theories could explain all aspects of the perception of a stable (or moving) world during eye movements. Direct theory could not explain the illusory motion of the visual world that occurs when a person presses against the eyeball or becomes dizzy (Helmholtz, 1910). Inferential theory required a veridical extra-

retinal, eye-velocity signal and it appears that proprioceptive or efferent signals usually underrepresent eye velocity (Filehne, 1922; Fleischl, 1882; Mack & Herman, 1973, 1978; Wertheim, 1987). More recent theories have incorporated aspects of both theories and hence have been better able to explain the phenomenon of stable perception during eye movements (e.g. Matin, 1982; Wallach, 1982; Wertheim, 1990).

The relevance to our observations is clear: if proprioceptive/efferent signals do in fact underrepresent eye velocity, those signals by themselves cannot provide an accurate estimate of the rotational flow component that must be cancelled in the flow field. Thus, the simple extra-retinal model schematized in Fig. 2 may not be an adequate explanation of how humans estimate heading during eye movements. If so, the visual system may use retinal-image-based schemes in conjunction with extra-retinal signals in estimating direction of self-motion. This possibility is discussed in more detail later in this section.

Because of its importance for the interpretation of our data, let us briefly review the experimental evidence for underrepresented eye velocity. Much of the evidence comes from experiments in which observers are asked to set the motion of a background stimulus to zero, in exo-centric coordinates, while making eye movements. When the background stimulus is a single line or dot in a dark field, or when it is presented briefly, observers typically set the velocity of the background in the same direction as and at a speed of about 20% of the eye movement; the gain of the observers' response is, therefore, roughly 0.8 (Mack & Herman, 1978; Wertheim, 1987). The response gain increases to about 1.0 when the background stimulus is large and textured and is presented for 1 sec or more (Mack & Herman, 1973; Wertheim, 1981, 1987). Presumably, the lower response gain with impoverished backgrounds provides a more accurate index of the gain of the extra-retinal, eye-velocity signal because the presence of a large, textured background allows retinal-image-based solutions, such as proposed by Gibson (1966, 1968),* to come into play. Similar effects are observed when observers report the perceived velocity of a moving object which they track with pursuit movements (Fleischl, 1882).† Thus, the ability to perceive a textured background's motion veridically while making smooth eye movements may well depend on contributions from extra-retinal signals and retinal-image-based algorithms for estimating object velocity in exo-centric coordinates.

In sum, there is persuasive evidence that extra-retinal signals underrepresent eye velocity during tracking eye movements. Thus, the extra-retinal model, as schematized in Fig. 2, may well be incorrect. Given the limitations of the extra-retinal model, a more robust means for estimating heading would use extra-retinal and retinal-image schemes and combine their outputs in some weighted fashion (perhaps with the weights varying according to viewing condition) in making the final heading estimate.

*In such cases, proponents of direct theory argue that the visual system assumes that large textured fields are stationary in exo-centric coordinates, so uniform retinal motion of such fields must result from an eye movement.

†Hansen (1979) examined peoples' spatial motor responses while tracking a moving spot in the dark. He found that people could strike a second, briefly flashed spot quite accurately with a hammer. Observers were not given feedback concerning the accuracy of their hammer strikes, so the accuracy of the responses indicates that the observers had rather accurate information about the position and velocity of the eyes in body-centered coordinates. This result is puzzling given the numerous perceptual experiments suggesting that extra-retinal signals underrepresent eye velocity. Further research is required to determine the source of the differences.

The misperception of heading during simulated rotations

In the simulated eye movement condition, heading judgments were displaced in the direction of the rotation by an amount roughly proportional to rotation rate. In this section, we examine two possible explanations for these consistent, but erroneous judgments.

Perhaps displaced estimates are a necessary consequence of estimating the translational component of the noisy flow fields during rotations. This hypothesis can be evaluated by examining the behavior of a model, such as Koenderink and van Doorn's (1987), that utilizes all the information in optic flow displays to estimate direction of self-motion. Koenderink and van Doorn's model, which in our terminology is a retinal-image model, calculated least-squares solutions to the flow equations (1) for flow fields derived from combinations of translations and rotations plus added noise. The model's performance depended critically on several parameters. For instance, as rotation rate increased beyond a critical value, the root mean squared (RMS) error of the model's heading estimates increased roughly in proportion to rotation rate. Apparently, the change in RMS error resulted from an increase in the variability of the estimates rather than from a change in the mean error. In contrast, human observers in our experiments exhibited mean errors proportional to rotation rate with little change in the variable error (Figs 7, 11, 13, and 15). Thus, human behavior in the presence of simulated rotations differs markedly from that of a model using all the information in the display efficiently. The most obvious cause of this difference is the extra-retinal information in the human that an eye rotation has not occurred. The visual system might seek solutions consistent with this information by, for example, assuming a rotation of zero and finding the translational component that leads to the best fit to the observed optic flow.

In the simulated eye movement condition, observers frequently perceived a curvilinear path of self-motion.* This erroneous percept stimulates a second hypothesis for explaining the displaced heading estimates in the simulated condition: perhaps the visual system misinterprets the stimulus as motion on a curved path and observers chose a point on the path for their heading response.

Why would observers perceive curvilinear motion in these displays? Movement of an observer on a circular path can be described *instantaneously* as the sum of a translation along the tangent to the path and an eye rotation about a perpendicular axis through the eye (Warren *et al.*, 1991). Consequently, the instantaneous flow field for forward translation plus an eye rotation about a vertical axis is identical to the field for curvilinear motion about a vertical axis with eye position fixed;

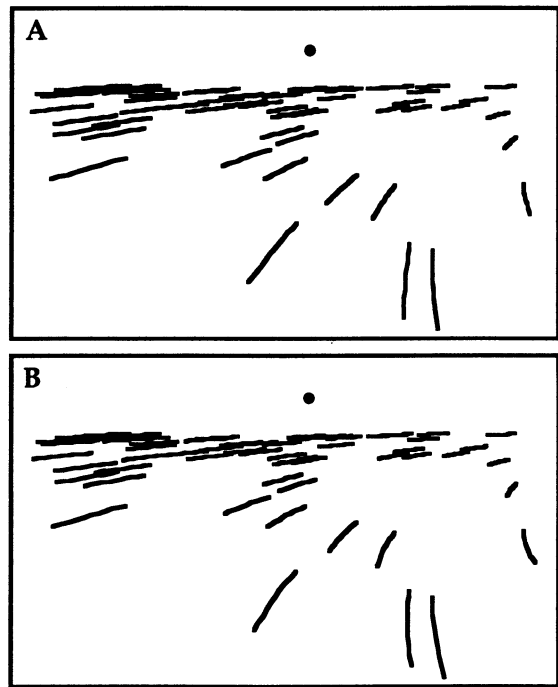


FIGURE 16. Retinal flow fields associated with linear translation plus rotation and curvilinear translation. (A) Trajectories of image points for forward translation across a ground plane while making a rightward eye movement. Translation speed = 190 cm/sec; rotation rate = 2.5 deg/sec; eye height = 160 cm; duration = 1 sec. (B) Trajectories of image points for curvilinear translation across a ground plane. Tangent of translation speed = 190 cm/sec; radius of curvature = 4354 cm; eye height = 160 cm; duration = 1 sec.

obviously, the two situations cannot be distinguished from the instantaneous fields. However, the sets of retinal image motions induced by the two situations differ more and more as time passes. To appreciate this, consider the two situations in terms of coordinate systems. For forward translation while making a horizontal eye movement (the situation in the simulated eye movement condition), the direction of translation does not change over time with respect to an exo-centric coordinate system but does change with respect to retino-centric coordinates. For curvilinear motion with eye position fixed (relative to the head and body), the direction of self-motion changes with respect to exo-centric coordinates but does not change relative to retino-centric coordinates.

Figure 16 illustrates the similarities and dissimilarities between the retinal image motions created by these two types of observer motion. Figure 16(A) displays the retinal trajectories of texture elements induced by forward translation across a ground plane plus an eye rotation about a vertical axis. Figure 16(B) displays trajectories induced by motion along a circular path with constant eye position. The sets of trajectories are quite similar, the greatest differences occurring for elements near the observer's feet. From the standpoint of retinal-image models, the visual system must distinguish these two sets of retinal trajectories in order to estimate heading accurately; the two may not be sufficiently different to do so. From the standpoint of extra-retinal models, the visual system can use the additional

*The percept of curvilinear self-motion through a rigid scene was most distinct when the stimulus consisted of one plane (Expts 4-6); some observers perceived curvilinear translation through a somewhat non-rigid scene when the stimulus consisted of two planes (Expts 1 and 2) or a three-dimensional cloud (Expts 7 and 8).

information that the eyes moved in the situation depicted in Fig. 16(A) and not in the one displayed in Fig. 16(B) and make the distinction on that basis. During simulated eye rotations, extra-retinal signals provide information that the eye has not moved, so the visual system may attempt to find observer motion parameters that are most consistent with an eye rotation of zero (in head-centered coordinates). Figure 16(A, B) illustrates that the observed flow is quite similar to that induced by curvilinear motion of the observer, so a reasonable reconciliation of the observed flow and the extra-retinal signal is to assume that the observer has moved on a curved path with zero rotation (in head-centered coordinates).^{*} This possibility is explored quantitatively elsewhere (Royden, 1994).

CONCLUSION

We examined the use of retinal-image and extra-retinal information in determining the direction of self-motion during eye movements. The critical comparisons involved judgments of heading with executed as opposed to simulated eye movements. In general, observers judged heading much more accurately during executed rather than simulated eye movements implying that extra-retinal, eye-velocity signals are used in determining the direction of self-motion. There were some experimental conditions in which observers could judge heading reasonably accurately during simulated eye movements; these included conditions in which eye movement velocities were 1 deg/sec or less and conditions which made available a horizon cue that exists for locomotion parallel to a ground plane. Thus, there appear to be some situations in which extra-retinal eye-velocity signals are not critically involved in estimating heading.

Overall, our results imply that extra-retinal, eye-velocity signals are used in determining direction of self-motion under many, perhaps most, viewing conditions. When those signals imply that the eye has not moved, the visual system misinterprets the translational plus rotational flow field as the consequence of a displaced or curvilinear path of self-motion.

REFERENCES

- van den Berg, A. V. (1992). Robustness of perception of heading from optic flow. *Vision Research*, *32*, 1285–1296.
- van den Berg, A. V. (1993). Perception of heading. *Nature*, *365*, 497–498.
- Bingham, G. P. (1993). Optic flow from eye movement with head immobilized: "Ocular occlusion" beyond the nose. *Vision Research*, *33*, 777–790.
- Bruss, A. R. & Horn, B. K. P. (1983). Passive navigation. *Computer Vision, Graphics, and Image Processing*, *21*, 3–20.
- Carme, T. (1992). *Computer vision: Theory and industrial applications*. New York: Springer.
- Chapman, S. (1968). Catching a baseball. *American Journal of Physics*, *36*, 868–870.
- Clocksink, W. F. (1980). Perception of surface slant and edge labels from optical flow: A computational analysis. *Perception*, *9*, 253–267.
- Crowell, J. A. & Banks, M. S. (1993). Perceiving heading with different retinal regions and types of optic flow. *Perception & Psychophysics*, *53*, 325–337.
- Crowell, J. A., Royden, C. S., Sekuler, A. B., Swenson, K. & Banks, M. S. (1990). Optic flow and heading judgments. *Investigative Ophthalmology and Visual Science (Suppl.)*, *31*, 522.
- Cutting, J. E. (1986). *Perception with an eye to motion*. Cambridge, Mass.: MIT Press.
- Droulez, J. & Cornilleau-Peres, V. (1990). Visual perception of surface curvature. The spin variation and its physiological implications. *Biological Cybernetics*, *62*, 211–224.
- Duffy, C. J. & Wurtz, R. H. (1991). Sensitivity of MST neurons to optic flow stimuli. I. A continuum of response selectivity to large-field stimuli. *Journal of Neurophysiology*, *65*, 1329–1345.
- Filehne, W. (1922). Ueber das optische Wahrnehmen von Bewegungen. *Zeitschrift für Sinnesphysiologie*, *53*, 134–145.
- Fleischl, E. (1882). Physiologisch-optische Notizen. *Sitzungsberichte der Akademie der Wissenschaften*, *86*, 17–25.
- Geisler, W. S. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological Reviews*, *96*, 267–314.
- Gibson, J. J. (1950). *The perception of the visual world*. Boston, Mass.: Houghton Mifflin.
- Gibson, J. J. (1966). *The senses considered as perceptual systems*. Boston, Mass.: Houghton Mifflin.
- Gibson, J. J. (1968). What gives rise to the perception of motion? *Psychological Review*, *75*, 335–345.
- Gibson, J. J., Olum, P. & Rosenblatt, F. (1955). Parallax and perspective during aircraft landings. *American Journal of Psychology*, *68*, 373–385.
- Gordon, D. A. (1965). Static and dynamic visual fields in human space perception. *Journal of the Optical Society of America*, *55*, 1296–1303.
- Hansen, R. M. (1979). Spatial localization during pursuit eye movements. *Vision Research*, *19*, 1213–1221.
- Heeger, D. J. & Jepson, A. D. (1992). Subspace methods for recovering rigid motion I: Algorithm and implementation. *International Journal of Computer Vision*, *7*, 95–117.
- Helmholtz, H. (1910). *Handbuch der physiologischen Optik* (Vol. 3). Leipzig: Voss.
- Hildreth, E. C. (1992). Recovering heading for visually-guided navigation. *Vision Research*, *32*, 1177–1192.
- von Holst, E. (1954). Relations between the central nervous system and the peripheral organs. *British Journal of Animal Behavior*, *2*, 89–94.
- von Holst, E. & Mittelstaedt, H. (1950). Das Reafferenzprinzip. *Naturwissenschaften*, *37*, 464–476.
- Howard, I. P. (1982). *Human visual orientation*. New York: Wiley.
- Koenderink, J. J. & van Doorn, A. J. (1976). Local structure of movement parallax of the plane. *Journal of the Optical Society of America*, *66*, 717–723.
- Koenderink, J. J. & van Doorn, A. J. (1981). Exterospic component of the motion parallaxfield. *Journal of the Optical Society of America*, *71*, 953–957.
- Koenderink, J. J. & van Doorn, A. J. (1987). Facts on optic flow. *Biological Cybernetics*, *56*, 247–254.
- Lee, D. N. (1980). The optic flow field: The foundation of vision. *Philosophical Transactions of the Royal Society of London B*, *290*, 169–179.
- Lisberger, S. G., Morris, E. J. & Tychsen, L. (1987). Visual motion processing and sensory-motor integration for smooth pursuit eye movements. *Annual Reviews of Neuroscience*, *10*, 97–129.
- Longuet-Higgins, H. C. (1981). A computer algorithm for reconstructing a scene from two projections. *Nature*, *293*, 133–135.
- Longuet-Higgins, H. C. (1984). The visual ambiguity of a moving plane. *Proceedings of the Royal Society of London B*, *223*, 165–175.
- Longuet-Higgins, H. C. & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London B*, *208*, 385–397.

^{*}There may be viewing conditions for which observers can distinguish linear translation plus rotation from curvilinear translation on the basis of retinal information alone. Such conditions might include a large field of view and long stimulus duration (Warren *et al.*, 1991). This remains to be examined experimentally.

- Mack, A. & Herman, E. (1973). Position constancy during pursuit eye movement: An investigation of the Filehne illusion. *Quarterly Journal of Experimental Psychology*, *25*, 71–84.
- Mack, A. & Herman, E. (1978). Loss of position constancy during pursuit eye movements. *Vision Research*, *18*, 55–62.
- Matin, L. (1982). Visual localization and eye movements. In Wertheim, A. H., Wagenaar, W. A. & Leibowitz, H. W. (Eds), *Tutorials on motion perception*. New York: Plenum Press.
- McLeod, P. & Denes, Z. (1993). Running to catch a flyball. *Nature*, *362*, 23.
- Perrone, J. A. (1992). Model for the computation of self-motion in biological systems. *Journal of the Optical Society of America A*, *9*, 177–194.
- Perrone, J. A. & Stone, L. S. (1994). The perception of egomotion: Global vs local mechanisms. *Nature*. Submitted.
- Regan, D. & Beverley, K. I. (1982). How do we avoid confounding the direction we are looking and the direction we are moving? *Science*, *215*, 194–196.
- Rieger, J. H. & Lawton, D. T. (1985). Processing differential image motion. *Journal of the Optical Society of America A*, *2*, 354–360.
- Rieger, J. H. & Toet, L. (1985). Human visual navigation in the presence of 3D rotations. *Biological Cybernetics*, *52*, 377–381.
- Royden, C. S. (1994) Analysis of misperceived observer motion during simulated eye rotations. *Vision Research*, *34*, 3215–3222.
- Royden, C. S., Banks, M. S. & Crowell, J. A. (1992). The perception of heading during eye movements. *Nature*, *360*, 583–585.
- Saito, H., Yukie, M., Tanaka, K., Hikosaka, K., Fukada, Y. & Iwai, E. (1986). Integration of direction signals of image motion in the superior temporal sulcus of the macaque monkey. *Journal of Neuroscience*, *6*, 145–157.
- Sperry, R. W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology*, *43*, 482–489.
- Wallach, H. (1982). Eye movement and motion perception. In Wertheim, A. H., Wagenaar, W. A. & Leibowitz, H. W. (Eds), *Tutorials on motion perception*. New York: Plenum Press.
- Warren, W. H. & Hannon, D. J. (1988). Direction of self-motion is perceived from optical flow. *Nature*, *336*, 162–163.
- Warren, W. H. & Hannon, D. J. (1990). Eye movements and optical flow. *Journal of the Optical Society of America A*, *7*, 160–169.
- Warren, W. H., Mestre, D. R., Blackwell, A. W. & Morris, M. W. (1991). Perception of circular heading from optical flow. *Journal of Experimental Psychology: Human Perception and Performance*, *17*, 28–43.
- Waxman, A. M. & Ullman, S. (1985). Surface structure and 3D motion from image flow: A kinematic analysis. *International Journal of Robotics Research*, *4*, 72–94.
- Wertheim, A. H. (1981). On the relativity of perceived motion. *Acta Psychologica*, *48*, 97–110.
- Wertheim, A. H. (1987). Retinal and extraretinal information in movement perception: How to invert the Filehne illusion. *Perception*, *16*, 289–294.
- Wertheim, A. H. (1990). In Warren, R. & Wertheim, A. H. (Eds), *Perception and control of self-motion*. Hillsdale, N.J.: Erlbaum.

Acknowledgements—The authors thank Kirk Swenson for software assistance, Rowan Candy and Irene Laudeman for their participation as observers, and Ellen Hildreth for helpful discussion. This research was supported in part by NIH Research Grant HD-19927, NSF Research Grants DBS-9309820 and IRI-9196196, and NASA-Ames Research Center Grant NAG 2-652. CSR was supported by a Life Sciences Research Foundation Fellowship from the Burroughs Wellcome Fund and a Science Scholar's Fellowship from the Bunting Institute of Radcliffe College. JAC was supported by Training Grant EY-07043.