Why two eyes are better than one for judgements of heading

A. V. van den Berg & E. Brenner

Physiology I, Medical Faculty, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands

ARE two eyes needed for judging direction of self-motion? Traditional analyses stress that the pattern of optic flow in one eye is sufficient[1-5]. The main difficulty is how to deal with the eye or head rotation. Extraretinal signals help[6-8], but humans can also discount the effect of rotation purely on the basis of monocular flow[6,7,9-12] provided the scene contains depth[6,9,10]. Depth differences give rise to changing binocular disparities when the observer moves. These disparities are ignored in monocular theories of judgements of heading. Using computer generated displays, we investigated whether stereoscopic presentation improves heading judgements for conditions that pose problems to the monocular observer. We found that adding disparities to simulated ego-motion through a cloud of dots made heading judgements up to four times more tolerant to motion noise. The same improvement was found when the disparities specify the initial distances throughout the motion sequence. We conclude that binocular disparities improve judgements of heading by imposing a depth order on the elements of the scene, not because they provide additional information on the elements' motion in depth.

When a driver fixates a mountain ridge in the distance, his direction of gaze is practically stationary, and the retina receives a motion pattern that radiates outward from his direction of heading. In contrast, when rotating his eye and his head so as to fixate a road sign, he will null the sign's motion on the fovea. In this case, the retinal motion pattern (retinal flow) radiates outward from the fixation point rather than from the destination point. How can humans disregard the rotational component, which complicates the judgement of heading? Normally, visual and extraretinal signals that accompany the self rotation work in concert to discount the rotation[6-8]. Nevertheless, when one presents the retinal motion of a rotating and translating observer to a stationary eye, heading is often perceived accurately[6,7,9-12]. Under such conditions, monocular heading judgements are sensitive to the layout of the environment. They are accurate in the presence of noise[7,12] or fast eye rotations[11] when motion across the ground plane is simulated, but not for motion through a cloud of dots[7,8]. Depth cues (perspective, texture gradients and height in the display) help to derive the heading from the retinal flow in the case of the ground plane[12]. Hence, we surmised that adding stereoscopic information to the flow would improve the performance for motion through the cloud.

We investigated the sensitivity of heading judgements to noise. In the first experiment we used presentations with and without stereoscopic information, both for the ground plane and for a cloud of dots. Horizontal simulated self-motion was always presented to both eyes, but in the synoptic (monocular information) case the eyes received identical images. We simulated the version and, in the stereoscopic condition, the vergence eye movements that were required to fixate a point in the environment. At the end of the motion sequence, the subject used a pointer to indicate the perceived direction of heading (Fig. 1).

For stereo presentations we found similar performance for simulated motion across the plane and for motion through the cloud (see example in Fig. 2). For motion across the ground plane, we found little difference between stereoscopic and synoptic presentation. In both cases, pointing was accurate and precise when the speed of the local motion vectors in the flow was four times larger than the speed of the local noise (signal-to-noise ratio = 4, see Fig. 1 legend). For lower SNR, precision decreased and subjects showed an increasing tendency to point towards the fixation point. Little correlation remained between the pointing responses and the simulated heading direction when the noise exceeded the signal (SNR < 1.0).

For motion through a cloud of dots, we found a clear difference between stereoscopic and synoptic presentations (circles in Fig. 3). For stereoscopic presentation heading was perceived

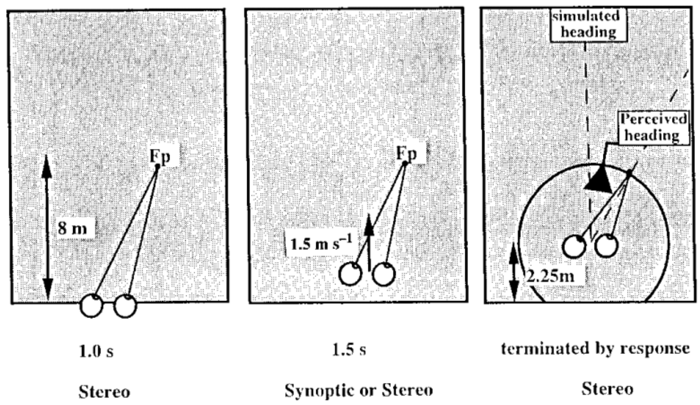

FIG. 1 The sequence of events during a trial. The scene contained about 256 white dots randomly distributed in a cloud or on the ground plane. The simulated depth range was from 1 to 20 m. Other dimensions were determined by the screen size (60° horizontally × 50° vertically). Subjects fixated a red point, that was part of the scene, at variable eccentricity and initially at 8 m distance. The simulated eye height above the ground plane was 0.65 m. To aid fixation, the first frame was shown stereoscopically for one second. Subsequently, forward motion was simulated with a speed of 1.5 m s^{-1}. During this period presentation was either synoptic or stereoscopic. The synoptic motion sequence corresponded to the motion pattern that would be received by a point at the bridge of the nose. Dot lifetime was limited to 160 ms to rule out the use of cues related to the trajectories of individual dots. Each dot's motion was perturbed with randomly directed noise. The magnitude of the noise component was proportional to the local flow velocity (SNR = v_{flow}/v_{noise}). The eye rotations required to fixate the red point (Fp) were simulated. Thus, the images of the red point for the two eyes were stationary on the screen. This imposed a fixed eye vergence that corresponded to a distance halfway between the initial and the final simulated positions of the red point. After 1.5 s the motion stopped. The scene was shown stereoscopically, with a triangular pointer, which the subject turned about a circle concentric with his feet so as to indicate the perceived direction of heading. A button press terminated the presentation. Subjects were told that the displays mimicked the view one would receive when looking at a road sign while driving a car. They were asked to indicate the heading direction of the car. All three subjects were given feedback on their performance during 10–50 training trials, but not during testing.

down to SNR = 1. For synoptic presentation, subjects could tolerate less noise and perceived heading down to SNR = 2 or SNR = 4. Note that as more noise was introduced (lower SNR), the precision decreased and subjects' responses became more biased towards the fixation point (Fig. 3).

In the second experiment we investigated whether changing disparity was essential for the improved performance in the cloud. The motion sequence was identical in the two eyes, but each dot of the cloud was given a fixed disparity that corresponded to the dot's simulated three-dimensional position in the first frame. Performance for this 'static-stereo' presentation was very similar to that for the full-stereo condition (unfilled symbols in Fig. 3). Thus, the depth order that static disparities impose on the dots in the cloud is sufficient to enhance performance to the level attained for the ground plane.

We have shown that stereoscopic depth is beneficial to the perception of heading. This is the case, not because it provides a direct cue to the motion in depth of the rigid environment relative to the observer, but because it provides a depth order, as could occlusion or texture gradients. The relative magnitude of the translatory and the rotatory contributions to the flow-field changes with the distance. We think that the independent information on the depth order allows the brain to exploit this property of the flow-field. Specifically, the most distant points are most reliable for estimation of the self-rotation, because the translatory part of the flow is inversely proportional to the distance. Conversely, for relatively small rotation rates the flow of nearby points is dominated by the translatory part. Such relations may be used to constrain the set of possible heading directions and ego-rotations that are consistent with the level, correlation between pointing responses and simulated heading was less than 0.5. The lower panels show the slopes of the perceived versus the simulated heading. Values lower than one indicate a bias towards the fixation point.

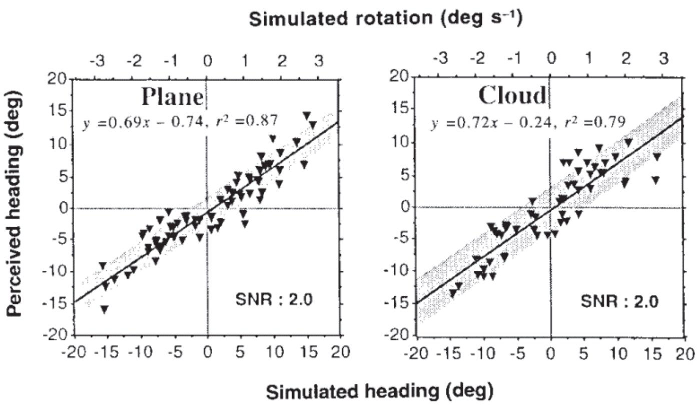

FIG. 2 Example of pointing responses for simulated motion across the plane and for motion through the cloud. Each point indicates the response in a single trial. Heading is expressed as an angle relative to the fixation direction at the end of the presentation. Perceived and simulated heading directions are linearly related. If the correlation exceeds a criterion level (0.5), we characterize the constant and the variable error of the subject's pointing response by the slope of the regression line and the s.d. of the perceived heading relative to the predicted heading (s.d.(ε)) using the regression line. Perfect pointing would correspond to a slope of 1.0 and negligible variation of pointing (s.d.(ε) = 0). This subject's (J.K.) perceived heading was biased towards the fixation point by about 30%, irrespective of the layout of the points (the slope of the regression line is about 0.7 in both cases). His variable part of the pointing error was also very similar (cloud: 2.9°; plane: 2.2°). The shaded area in each panel indicates a range of 2 s.d.(ε) about the best fitting regression line. The upper axis indicates the average simulated horizontal rotation rate.

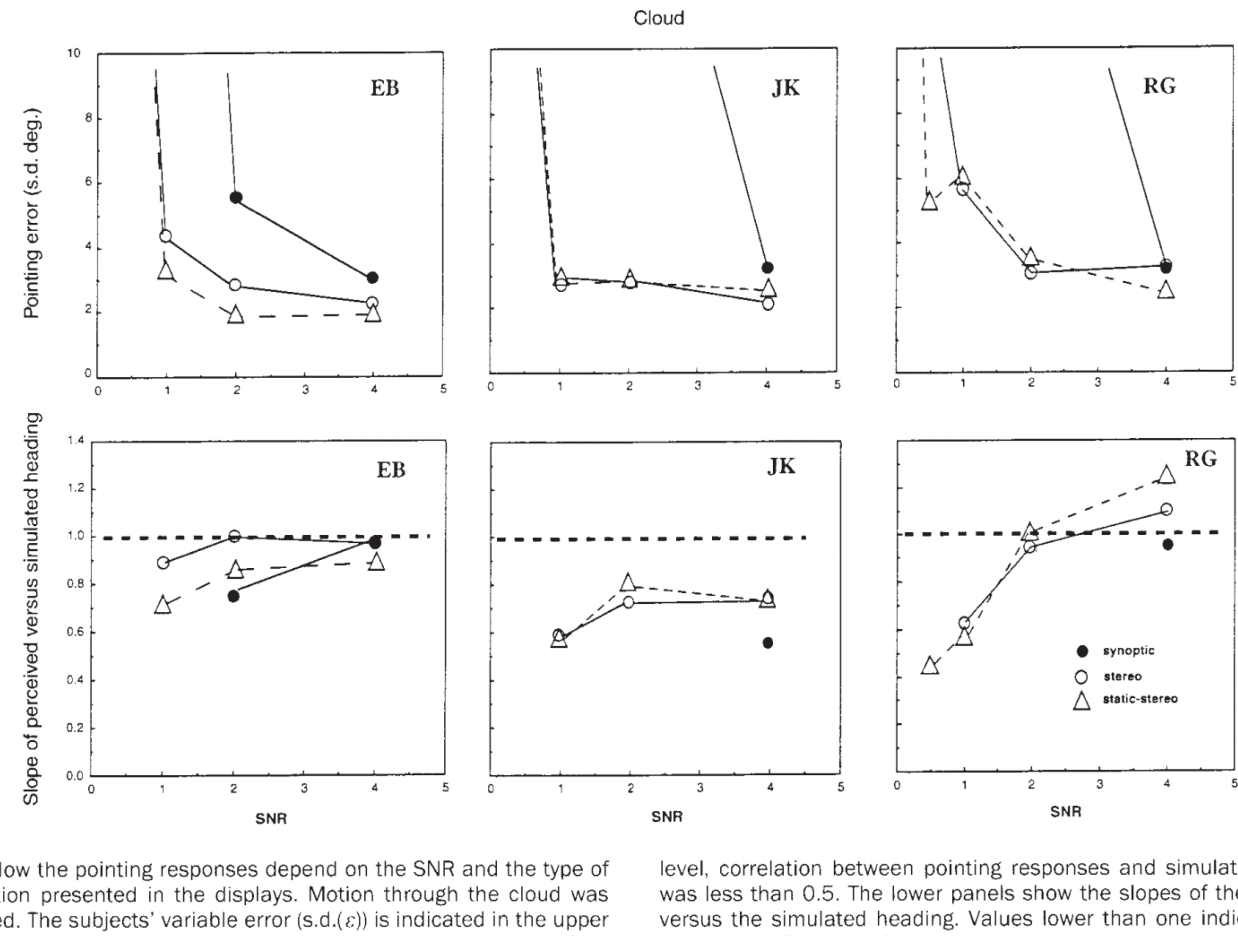

FIG. 3 How the pointing responses depend on the SNR and the type of information presented in the displays. Motion through the cloud was simulated. The subjects' variable error (s.d.(ε)) is indicated in the upper panels. The steep upward lines indicate that at the next lower SNR level, correlation between pointing responses and simulated heading was less than 0.5. The lower panels show the slopes of the perceived versus the simulated heading. Values lower than one indicate a bias towards the fixation point.

observed flow-field, resulting in reduced scatter in the perceived heading. Without static depth information, visual heading judgements are more vulnerable to noise and the confounding effects of eye and head rotation. □

Received 30 March; accepted 26 August 1994.

1. Gibson, J. J. *Perception of the Visual World* (Houghton Mifflin, Boston, 1950).
2. Longuett-Higgins, H. C. & Prazdny, K. *Proc. R. Soc.* B**208,** 385–397 (1980).
3. Rieger, J. H. & Lawton, D. T. *J. opt. Soc. Am.* A**2,** 354–359 (1985).
4. Koenderink, J. J. & Van Doorn, J. A. *Biol. Cybern.* **56,** 247–254 (1987).
5. Heeger, D. J. & Jepson, A. *Neural Comp.* **2,** 129–137 (1990).
6. Warren, W. H. & Hannon, D. J. *Nature* **336,** 162–163 (1988).
7. van den Berg, A. V. *Vision Res.* **32,** 1285–1296 (1992).
8. Royden, C. S., Banks, M. S. & Crowell, J. A. *Nature* **360,** 583–585 (1992).
9. Warren, W. H. & Hannon, D. J. *J. opt. Soc. Am.* A**7,** 160–169 (1990).
10. Rieger, J. H. & Toet, L. *Biol. Cybernet.* **52,** 377–381 (1985).
11. van den Berg, A. V. *Nature* **365,** 497–498 (1993).
12. van den Berg, A. V. & Brenner, E. *Vision Res.* **34,** 2153–2167 (1994).