Pergamon

PII: S0042-6989(96)00003-X

Vision Res., Vol. 36, No. 23, pp. 3805–3814, 1996
Copyright © 1996 Elsevier Science Ltd. All rights reserved
Printed in Great Britain
0042-6989/96 $15.00 + 0.00

The Special Role of Distant Structures in Perceived Object Velocity

ELI BRENNER,*† A.V. van den BERG*

Received 31 March 1995; in revised form 21 September 1995

How do we judge an object's velocity when we ourselves are moving? Subjects compared the velocity of a moving object before and during simulated ego-motion. The simulation consisted of moving the visible environment relative to the subject's eye in precisely the way that a static environment would move relative to the eye if the subject had moved. The ensuing motion of the background on the screen influenced the perceived target velocity. We found that the motion of the "most distant structure" largely determined the influence of the moving background. Relying on retinal motion relative to that of distant structures is usually a reliable method for accounting for rotations of the eye. It provides an estimate of the object's movement, relative to the observer. This strategy for judging object motion has the advantage that it does not require metric information on depth or detailed knowledge of one's own motion. Copyright © 1996 Elsevier Science Ltd

Optic flow Perspective Stereopsis Motion Eye movements

INTRODUCTION

How do we judge a visible object's velocity when we ourselves are moving? The most obvious possibility would be to make some kind of "prediction" (though not necessarily a conscious one) of how our movements would have shifted the object's retinal image if the object were stationary. The difference between the predicted and the actual retinal motion can then be attributed to motion of the object.

Knowing our own movements would help make such predictions. However, knowing our own motion is not enough. We also need to know the object's distance. To avoid confusion, we will use the terms eye-rotation and eye-translation to refer to the rotation and translation of our eyes *relative to the surroundings.* We do so to avoid the term "eye movements", which is used to describe the rotation of the eyes *relative to the head.* It is easy to predict how eye-rotation influences the object's retinal motion. Rotations shift the whole image on the retina to the same extent. In contrast, without knowing the object's distance, it is impossible to predict how eye-translation shifts the object's retinal image. Translation shifts the images of structures in the environment in inverse proportion to their distances from the eye. To predict the object's retinal motion, therefore, requires independent information on the object's distance.

If we are moving through a rigid, stationary environment, the changing perspective as a result of our motion gives rise to systematic changes in the image of the environment on our retina (Gibson, 1979; Koenderink, 1986). These systematic changes (the optic flow) provide us with information on the structure of the surroundings (Rogers & Graham, 1979; Cornilleau-Pérès & Droulez, 1994) as well as on our own motion (Warren & Hannon, 1988; Van den Berg, 1992). Additional information on our own motion is normally available from various extra-retinal sources, such as vestibular stimulation, proprioception, and so on (e.g. Mergner *et al.*, 1992). Similarly, extra-retinal information on the orientation of our eyes can help us localise the object when we fixate it.

In the present paper we will concentrate on target motion in the frontal plane and lateral eye-translations (parallel to the target's trajectory). When subjects only have *extra-retinal* information on their own motion (i.e., when they make real lateral movements in the dark), the target distance specified by ocular convergence influences the perceived target motion (though clearly not to the extent that would be required for accounting for one's own movements; Gogel, 1982; Schwarz *et al.*, 1989). In contrast, when subjects only have *retinal* information on their own motion (i.e., when ego-motion is simulated by moving the environment), the target distance specified by ocular convergence (and relative disparity) does not influence the perceived target motion (Brenner, 1991). The latter finding cannot be due to the simulation having been interpreted as motion of the environment (which it actually was) rather than as ego-motion, because the moving environment did influence the perceived target

*Department of Physiology, Erasmus University, P.O. Box 1738, 3000 DR, Rotterdam, The Netherlands.
†To whom all correspondence should be addressed [*Fax* 31 10 4367594; *Email* brenner@fys1.fgg.eur.nl].

velocity considerably. A possible explanation is that the visual information that is used to judge an object's velocity while one is moving does not consist of separate judgements of ego-motion and target distance, but of aspects of the image that provide direct estimates of the object's motion. In this study we consider two such possibilities.

If the object is moving across a surface that is part of the stationary environment, one could rely on local relative motion. The retinal image of the part of the surface that the object is moving across will undergo the same shift due to both eye-rotations and eye-translations as does the object itself, because they are at the same distance. In this way, local relative motion could provide judgements of an object's motion *relative to the surroundings*. However, if the structures that have retinal images adjacent to that of the object are not at the same distance from the observer as the object, the object's velocity will be misjudged.

Another way of judging object motion without using metric information on distance is by relying on retinal motion relative to that of the most distant structure. To do so, the observer has to determine which visible structure is furthest away from himself. This information could, for instance, be obtained from perspective. The retinal images of distant structures are hardly shifted by eye-translations. Eye-rotations shift them to the same extent as they do any other structure. The retinal motion of distant structures (when expressed as an angular velocity) therefore provides a direct estimate of the influence of eye-rotations on all retinal motion. This estimate is reliable as long as the distant structures are indeed far away (in terms of the velocity of eye-translation). We could, therefore, account for eye-rotations by judging all retinal velocities relative to that of the most distant structure available. In doing so, we would obtain an estimate of object motion *relative to our (translating) eye*; irrespective of changes in the orientation of the eye (eye-rotations). Although this is contrary to our intuitive impression of perceived motion, because it implies—for instance—that a stationary target will appear to move when we ourselves move, it could still be the basis for perceived motion, with the distinction as to what had actually moved (oneself or the object) deferred to a later stage. An implication of this option is that when the distant structures are not far away, object motion will be systematically misjudged during eye-translation.

In our previous study, in which we found no influence of target distance (Brenner, 1991), the target was at the centre of a distant, frontal plane, well above a simulated horizontal surface. Thus, the target's local surrounding was the most distant surface. The finding that the simulated distance had no influence is, therefore, consistent with judging target velocity both on the basis of local relative motion and on the basis of motion relative to the most distant structure. We previously presented some evidence that the results were unlikely to (only) be due to the use of local relative motion. In the present study we examine this in more detail, with an emphasis on whether subjects use the retinal motion of the most distant structure in the proposed manner when estimating a target's velocity.

EXPERIMENT 1

In the first experiment we examine whether modifying the stimulus so that the target's local surrounding is no longer the most distant surface influences the results. The stimulus was similar to that of the previous study, but the target moved across a horizontal, ground surface. When target distance was varied, the target's *angular* velocity remained the same. This was achieved by scaling the target's simulated velocity together with its simulated distance. The angular velocity of the most distant structure was independent of target distance, because the simulated ego-motion was always the same. Thus, judging motion relative to the most distant structure predicts the same results (when expressed as angular velocities) for all target distances, whereas local relative motion predicts different angular velocities for different distances, because the local angular velocity of the background varies across the scene.

Methods

The experiments were conducted using a Silicon Graphics GTX-210 Computer with an HL69SG monitor. The image on the screen was 34 cm wide (1280 pixels) and 27 cm from top to bottom (492 pixels). Subjects sat with their head in a chin-rest at 42 cm from the screen; resulting in an image of 44×36 deg of visual angle. Images were presented at a frame rate of 120 Hz. LCD shutter spectacles ensured that alternate frames were presented to the left and right eyes. Red stimuli were used because the LCD shutter spectacles work best at long wavelengths (about 33% transmission when "open" and 0.3% when "shut"). Screen luminance was 13 cd/m^2 for light pixels and 0.02 cd/m^2 for dark ones. Each image was drawn in appropriate perspective for the eye that saw it, and for the simulated positions of the target and the observer at that instant. Apart from the stimulus, the room was completely dark.

The display is shown schematically in Fig. 1. The target was a small cube that moved from left to right across a simulated horizontal plane. This simulated plane and a simulated, distant, frontal surface were covered with small squares. During the first part of each presentation, these two surfaces were static. Only the target moved. During the second part of each presentation, the two surfaces could move to the left (with the appropriate changes in perspective). We refer to this stimulus as a simulated eye-translation (to the right).

The target was simulated to either be halfway between the observer and the frontal surface, three-quarters of the distance to the surface, or immediately in front of it. The frontal surface, which was the most distant visible structure, was close enough for its image to move considerably on the screen during the simulated eye-translation (see legend of Fig. 2). Relying on motion relative to the most distant structure, in the manner

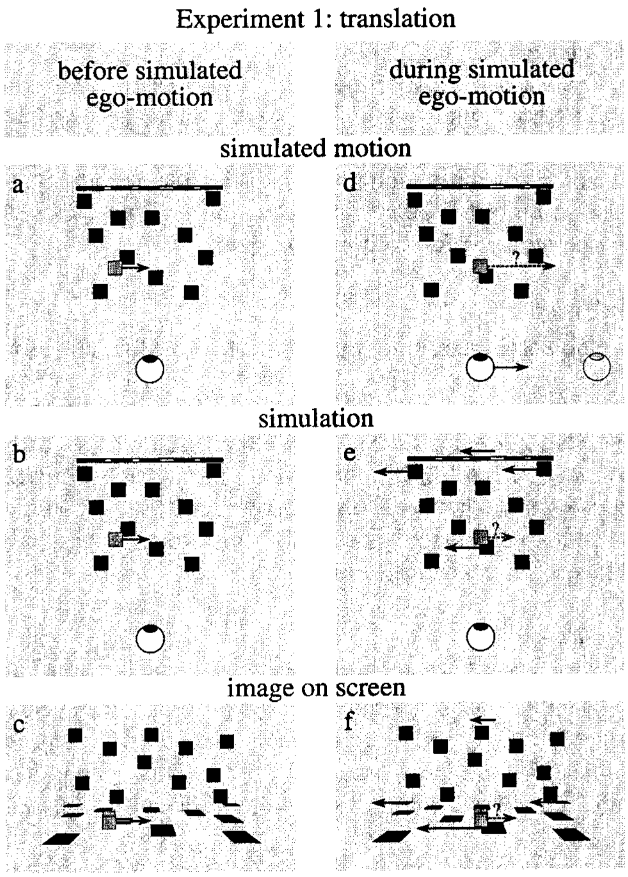

FIGURE 1. Schematic representation of the stimulus before and during simulated ego-translation. The simulation consisted of moving the background relative to the subject's static eye, in precisely the way that it would have moved relative to the eye if the eye had moved. The background consisted of 105 squares distributed at random on a horizontal and a frontal surface. Before the simulated ego-motion [seen from above in (a) and (b)], only the target moved on the screen [arrow in (c)]. Ego-motion [arrow at "eye" in (d)] was simulated by moving all the squares relative to the static eye [arrows in (e)]. The velocity with which each square moved across the screen [length of arrows in (f)] was inversely proportional to the surface's simulated distance, so that the squares nearby on the horizontal surface moved twice as fast as those at the back. We examined how fast the target had to move during the simulated ego-motion (dashed arrow) for it to appear to continue to move at the same velocity. Note that the simulation is a pure lateral translation of the eye [see thin outline in (d) for the eye's simulated position and orientation some time later].

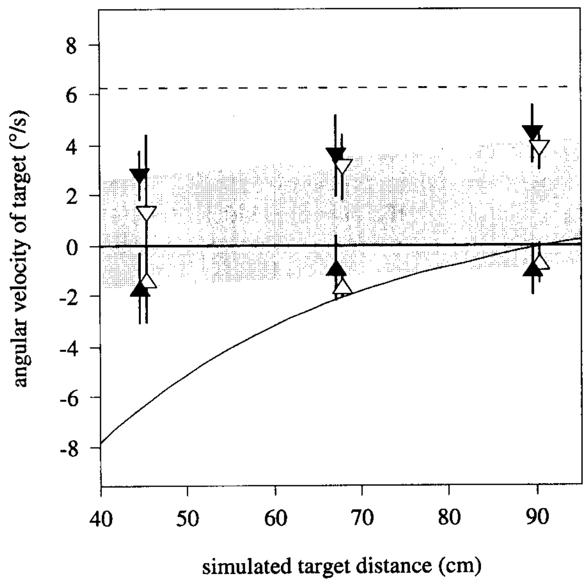

FIGURE 2. Range of angular velocities for which the object appeared to move at the same speed before and during simulated ego-motion (shaded area). Triangles pointing downwards and upwards are respectively the upper and lower limits of the range (average of five subjects with standard deviation between subjects). The target's initial angular velocity was slightly over 6 deg/sec to the right (dashed line). Simulated ego-motion at 10 cm/sec to the right shifts surfaces at 45 cm (distance of the nearest target) to the left at about 12 deg/sec and ones at 90 cm (distance of furthest target) to the left at about 6 deg/sec. In order to maintain the simulated target velocity, subjects would have to compensate for such shifts (thin curve). For targets moving across a surface, they could do so by maintaining the local relative velocity. The judged object velocity would be relative to the surroundings. In order to judge objects' velocities relative to themselves, subjects only have to account for their eye-rotations. If they use extra-retinal information to estimate their eye-rotation, they should simply maintain the target's angular velocity (dashed line). If they use the retinal slip of the image of the most distant structure to estimate eye-rotation, the movement of the background will be mistaken for the consequence of a rotation (thick line). The only proposal that falls within the experimentally determined range of subjective equality is that of judging object velocity relative to the most distant structure. The similarity between the data with (solid symbols) and without (open symbols) distance information from binocular stereopsis suggests that perspective determines which structure is considered the most distant.

proposed in the Introduction, would attribute this motion on the screen to eye-rotation, rather than translation, and, therefore, make subjects misjudge the target's velocity.

The cube initially moved at slightly more than 6 deg/sec. It filled about 1.3 deg of visual angle (both horizontally and vertically). The extent to which the cube's surfaces were visible depended on the cube's position and the distance between the observer's eyes. Images were calculated separately for each subject (and position), taking the distance between the individual subject's eyes into account. Apart from the differences in binocular cues, the nearby target was lower on the screen, and the image of its upper surface accounted for a larger part of its vertical dimension.

Both optic flow and perspective only provide distances relative to a scaling factor. The sizes, distances and velocities given below are all based on the assumption that subjects use the distance between their eyes as the scaling factor. This places the simulated horizontal surface 10 cm below the subject's eyes, and the distant frontal surface (50×20 cm) at a distance of 91 cm. This is the only scaling factor for which the relationships between distances specified by perspective and binocular stereopsis are consistent. However, if subjects do not use the distance between their eyes as the scaling factor, but, for instance, use their eye height instead (assuming that the horizontal plane is the ground they are standing on), all simulated sizes and velocities will be about 17 times larger. The angular velocity obviously does not depend on the scaling factor.

With the distance between the eyes as the scaling factor, the cube moved at simulated distances of 45, 67.5

or 90 cm from the observer. To be certain that we do not confound the influence of retinal size and angular velocity with that of simulated distance, the target's angular size and initial angular velocity were the same for all distances. As a result, the *simulated* target size and velocity changed in proportion to the simulated target distance: at 45 cm the cube had sides of 1 cm and moved at 5 cm/sec; at 67.5 cm this was 1.5 cm and 7.5 cm/sec; and at 90 cm it was 2 cm and 10 cm/sec. The frontal and horizontal planes contained no visible structures other than 35 (frontal plane) and 70 (horizontal plane) 2×2 cm squares.

The experiment was conducted under two conditions: with and without binocular stereopsis. In the former condition, the images presented to each eye corresponded with that eye's position, as described above. In the latter condition, images presented to both eyes were identical showing the view from a position midway between the eyes (alternate images were presented to the two eyes with the LCD spectacles, but this did not provide any additional depth information). The images were superimposed on the screen, so that the vergence angle required to fuse the images corresponded with the screen distance of 42 cm.

The way in which the velocities before and during simulated ego-motion were compared was essentially the same as in previous studies (Brenner, 1991, 1993; Brenner & van den Berg, 1994). Subjects were presented with a target moving to the right across a static background for between 500 and 750 msec (random variations in duration were used to discourage subjects from relying on the target's initial and final position, rather than on its velocity). When the target was at the vertical midline, the background suddenly started moving to the left, simulating rightward motion of the observer (at 10 cm/sec). At the same moment the target's velocity could change. The target and observer moved at their new velocities for another 250–500 msec, after which subjects had to indicate whether the target moved faster, at the same speed, or more slowly during the simulated ego-motion.

For finding the velocity at which subjects switched from seeing no change in velocity to seeing an increase in velocity, the staircase procedure was as follows: if the subject reported that the target accelerated, the target's final speed was set lower on the next presentation. If the subject either reported that it did not change its speed, or that its motion during the final interval was slower, its final speed was set higher on the next presentation. The magnitude of the increase or decrease was reduced (to 80% of the previous value) after each trial (in 11 steps from 5 to 0.5 cm/sec). The value onto which the staircase converged was taken as the upper limit of the range of subjective equality (the transition from no perceived change to a perceived increase in velocity). The lower limit of the range of subjective equality (transition from no perceived change to a perceived decrease in velocity) was determined in the same manner, except that reports of no change in speed resulted in a lower (rather than a higher) velocity on the next presentation (for additional details see Brenner, 1991).

The staircases for all distances, for presentations with and without binocular information on distance, and for both the upper and lower limits of the range of subjective equality, all ran simultaneously, with the specific staircase to be tested determined at random (from those not yet completed) for each presentation.

Subjects

Subjects were one of the authors (EB) and four colleagues who did not know the purpose of the experiment. The only instruction subjects received was that they should indicate whether the target moved faster, at the same speed, or more slowly during the simulated ego-motion. They were not instructed on what to do when in doubt, but had to choose one of the three responses. All subjects have normal binocular vision.

Results

Figure 2 shows the range of angular velocities during the simulated ego-motion for which the target appeared to continue to move at the same speed (the range of subjective equality; shaded area). This range was influenced by the target's simulated distance, but only slightly.

If subjects had ignored the background altogether, they would have required that the target more or less maintains its angular velocity relative to themselves for it to appear to continue moving at the same speed (thin dashed line). They did not. In fact they required a decrease in angular velocity that is close to the decrease that maintains the target's retinal motion relative to that of the most distant structure (thick line).

The conditions in the experiment were such that the decrease in angular velocity that maintains the retinal motion relative to that of the most distant structure was independent of the target's distance: the velocity of ego-motion (10 cm/sec) and the distance of the most distant structure (91 cm)—and thereby the most distant structure's angular velocity—were always the same. The actual required decrease in angular velocity (shaded area) does appear to depend slightly on the target distance, but this is much less than would be needed to maintain the local relative velocity (thin curve). Note that the conditions were favourable for relying on local relative motion: the target was small and seen slightly from above; the horizontal surface was quite densely structured; and the top of the cube was separated by almost 5 deg of visual angle from the bottom of the frontal surface when the cube was on its nearest path.

The results were extremely similar for binocular simulations (solid symbols) and for simulations in which binocular information specified that the image was flat (open symbols). This supports the notion, raised in the Introduction, that subjects use a strategy that does not require metric information on depth. After the experiment, subjects were asked whether they had experienced vection (that they themselves were moving) at any time during the experiment. None ever did. They saw the target move to the right and the surroundings to the left; in complete agreement with all extra-retinal information. Nevertheless, their judgements of object velocity were influenced by the simulated ego-motion. This too supports the use of aspects of the image that provide direct estimates of the object's motion, rather than separate judgements of ego-motion and target distance.

Discussion

It is evident from Fig. 2 that subjects do not maintain the simulated target velocity; neither relative to the environment (for instance by relying on the local relative velocity) nor relative to themselves (by ignoring the background). The perceived velocity was maintained when the relative motion between the target's retinal image and that of the distant frontal surface was identical before and during the simulated ego-motion. This suggests that the retinal motion of the image of the most distant surface is used to estimate the rotation of our eyes relative to the surroundings. As the axes of rotation are different for different parts of our body, and we seldom move only our eyes (e.g. Land, 1992), it may not always be feasible to obtain reliable extra-retinal predictions of the retinal motion caused by our rotations. The proposed mechanism only requires that we identify the most distant structure; presumably from the depth order provided by perspective. The resulting judgements of object velocity are relative to the observer's eye, disregarding changes in the eye's orientation.

One shortcoming of this experiment is that the outcome could also be interpreted as a compromise between local relative motion and absolute motion based on extra-retinal information. Such a compromise (often referred to as a low gain for the influence of background motion) has been found for various tasks (e.g. Raymond *et al.*, 1984; Post & Lott, 1990; Smeets & Brenner, 1995). Although we initially found that retinal information dominates the perceived velocity in this task (Brenner, 1991), we have since found that extra-retinal information can be quite important under some conditions (Brenner & van den Berg, 1994). A compromise between local relative motion and the actual angular velocity could also account for the (modest) effect of target distance in Fig. 2.

A second shortcoming of the first experiment is that the most distant structure is very large in terms of visual angle, so that the retinal motion of the most distant structure is also the most preponderous retinal motion. We therefore conducted a second experiment in which there was no frontal plane at the end of the horizontal plane, and the predicted direction of the effect was different for the two hypotheses proposed in the introduction.

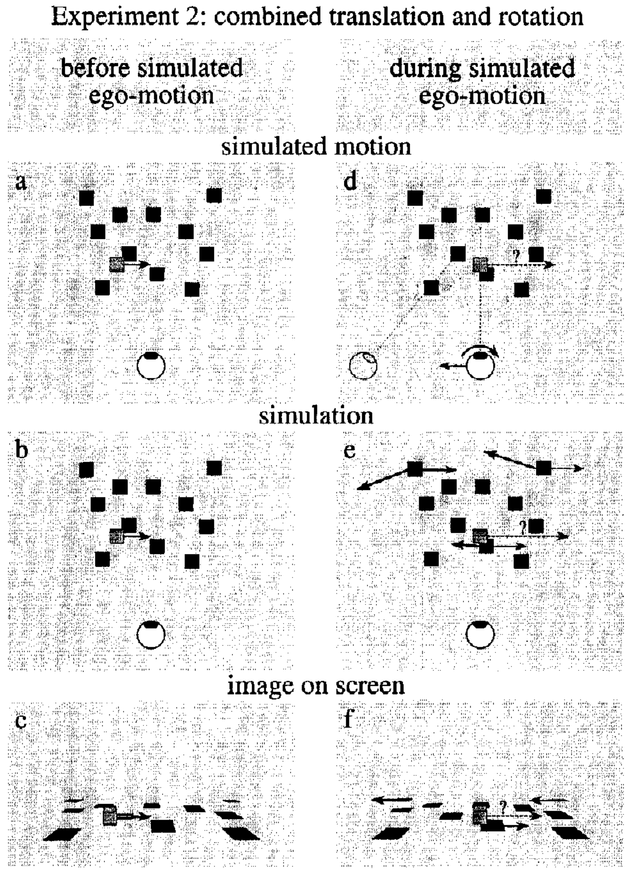

FIGURE 3. Schematic representation of the stimulus before and during a combination of simulated ego translation and rotation. The horizontal surface (represented by squares) actually consisted of 100 triangles. The rightward rotation and leftward translation [arrows in (d)] are simulated by moving the triangles in the appropriate manner relative to the observer [(e): thin arrows correspond with the simulated translation; thick arrows with the simulated rotation]. The influence of the simulated translation is larger than that of the simulated rotation for nearby structures, and smaller for distant structures, so that nearby structures move to the right whereas distant structures move to the left [on the screen; arrows in (f)]. The thin outline in (d) shows the eye's simulated position and orientation some time later. Note that the simulated ego-motion is a translation to the left while fixating a point behind the target.

EXPERIMENT 2

In order to have opposite directions of background motion for the most distant structures and for the structures closest to the target, we simulated a combination of ego rotation and translation. As the influence of translation depends on the simulated distance, whereas that of rotation does not, we can combine simulated rotation and translation in such a way that the most distant structures move to the left at the same velocity as the structures closest to the target move to the right. This corresponds with moving to the left while maintaining fixation on a point behind the target (see Fig. 3). The influence of this complex pattern of background motion was compared with that of uniform background motion (simulated ego-rotation; see Fig. 4). The most distant structures moved at the same velocity in both conditions (Table 1). In this experiment we also used a larger field of view, with the simulated floor coinciding with the real floor, in an attempt to make the simulation more "realistic".

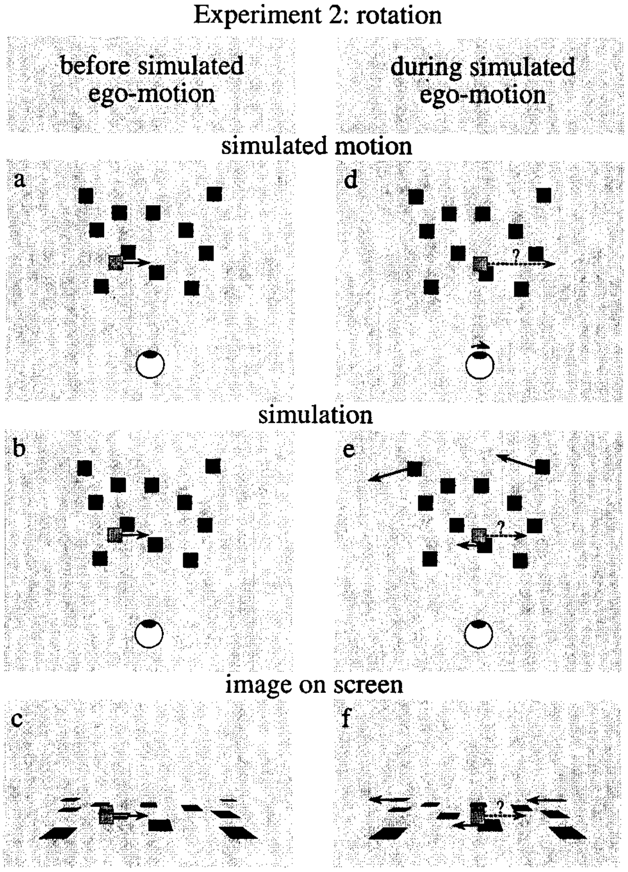

FIGURE 4. Schematic representation of the stimulus before and during simulated ego-rotation. The horizontal surface (represented by squares) actually consisted of 100 triangles. The rightward rotation (d) is simulated by moving the triangles in the appropriate manner relative to the observer (e). This results in coherent leftward motion of all the triangles on the screen (f).

Methods

The experiment was very similar to the first one. This time, the stimulus was projected onto a large screen using a Sony VPH-1270QM Multiscan Video Projector. The image on the screen was 174 cm wide (1280 pixels; 82 deg at 100 cm) and 188 cm from top to bottom (492 pixels; 86 deg at 100 cm; bottom 42 cm above the floor). Subjects stood with their backs against a frame at 100 cm from the screen. Images were back-projected onto the screen at a rate of 120 Hz. LCD shutter spectacles ensured that each frame could only be seen with one eye. The eye that was stimulated alternated between frames. Each frame provided a new image; calculated for that eye and simulated displacement. Different images were presented to the two eyes, taking account of the individual's inter-ocular distance.

The ground plane was simulated to correspond with the floor level, taking the individual subject's eye-height into account. One hundred randomly oriented triangles (with sides of 25 cm) were distributed in a semi-random fashion across the ground plane. Only these triangles were visible. Each triangle was first assigned a random distance lying between the closest position we could present on the screen (about 125 cm, depending on the subject's eye height) and the most distant position in our simulated environment, which we set at 600 cm. The triangle was then assigned a random lateral position within the range of positions that would be visible on the screen. This procedure was necessary to ensure that there were always structures on the ground surface in the vicinity of the target.

The target was a cube with sides of 20 cm. It always moved to the right, 100 cm behind the screen (200 cm from the subject). Its initial simulated velocity (before simulated ego-motion) was always 1 m/sec (thus its image moved at 50 cm/sec—about 27 deg/sec—across the screen). The target's velocity during the simulated ego-motion was varied as in the first experiment, with the step size decreasing from 0.5 to 0.01 m/sec.

There were nine conditions (see Table 1). The only difference between the conditions was the kind and speed of simulated ego-motion. There was one condition without any simulated ego-motion, four with simulated rotation (turning to the right at four different velocities) and four simulating a combination of translation to the left and rotation to the right. In the latter four conditions,

TABLE 1. The simulated ego-motion, and how the simulation influences the motion of selected parts of the background's image on the screen

	Simulated		Background velocity at	
	Rotation (deg/sec)	Translation (m/sec)	target distance (deg/sec)	largest distance (deg/sec)
Static observer	0	0	0	0
Simulated rotation	5	0	−5	−5
	9	0	−9	−9
	14	0	−14	−14
	18	0	−18	−18
Simulated rotation and translation	9	−0.5	5	−5
	18	−1	9	−9
	27	−1.5	14	−14
	33	−2	18	−18

For the simulated ego-motion, rotating rightwards (as when one pursues a target moving to the right) and translating to the right are considered positive. For the background, rightward motion of the image is considered positive. The target initially moved at about 27 deg/sec across the screen. The simulated target distance was 2 m. The largest simulated distance was 6 m.

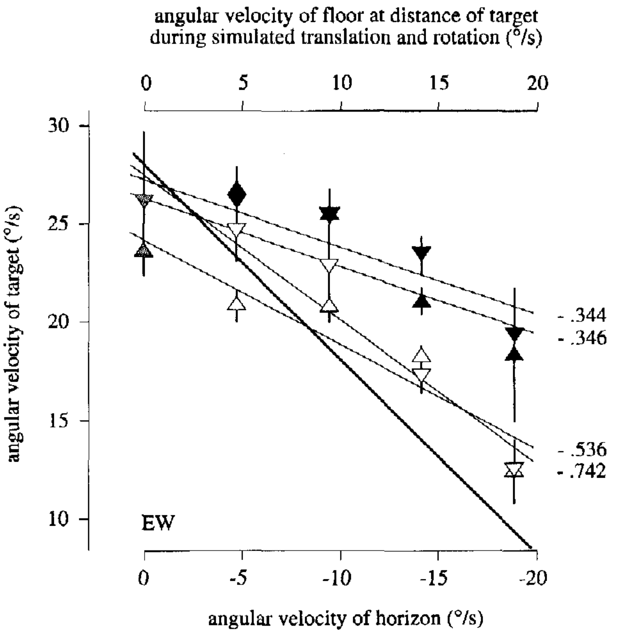

FIGURE 5. Means and standard errors of the outcomes of three replications for each of the nine conditions in experiment 2 (see Table 1) for one subject (EW). White triangles: simulated rotation. Black triangles: combined rotation and translation. Shaded triangles: no simulated ego-motion. Thick line: constant velocity relative to the horizon. Triangles pointing downwards and upwards are respectively the upper and lower limits of the range of simulated target velocities for which subjects reported perceiving no change in speed. The thin lines are the best fitting lines for the upper and lower limits for each kind of simulation. The numbers at the right are these lines' slopes. The averages of the two slopes for simulated rotation (in this case -0.536 and -0.742) and of the two slopes for combined rotation and translation (-0.344 and -0.346) are shown for each subject (with the slopes' average standard errors) in Fig. 6. The angular velocity of the background at the distance of the target (upper axis) is identical to that of the horizon (lower axis) during simulated rotation.

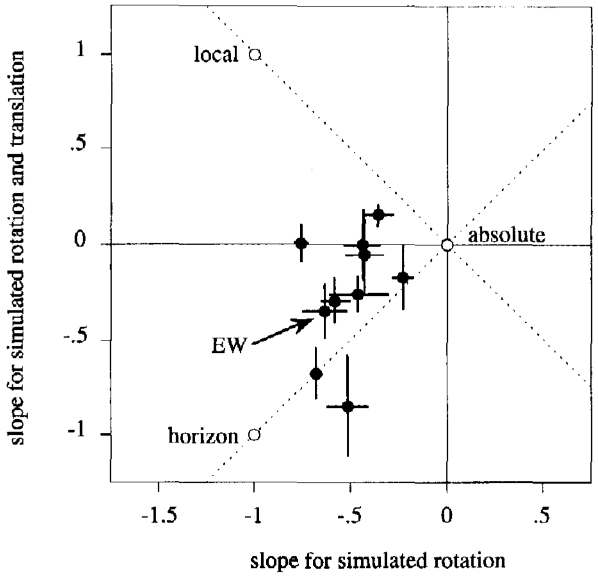

FIGURE 6. Influence of the background under the two kinds of simulations. Individual subjects' slopes for the change in angular velocity of the target (required to maintain the perceived velocity) as a function of the angular velocity of the horizon (see Fig. 5). The open symbols show where the points would be expected if one were to judge object motion exclusively in terms of the object's displacement relative to oneself (absolute), local relative motion (local), or motion relative to the most distant structure (horizon). Note that most subjects appear to base their judgements on a compromise between the absolute velocity and the velocity relative to the horizon.

the rates of simulated translation and rotation were calculated to result in the same leftward motion of the images of structures at 600 cm (the most distant structures) across the screen as for the simulated rotations, but with the structures at 200 cm (the target distance) moving at the same velocity in the opposite direction. All other aspects, such as the random interleaving of staircases, were as in the first experiment.

Subjects

Ten subjects took part in the second experiment. All except the two authors were unaware of the hypothesis being tested, but were aware of the fact that we were studying the role of background motion on perceived velocity. Four of the subjects performed the complete experiment three times, whereas the other six performed it once (the variability within subjects was considerably smaller than that between subjects). In contrast with our usual procedure, subjects were explicitly asked to only indicate that the target appeared to continue moving at the same speed when they were quite sure that this was so. We hoped that this explicit instruction would reduce the variability between subjects (which it did not).

Results

Figure 5 shows one subject's data for the two types of simulations. The average outcome of the staircases for each kind of simulated ego-motion are shown by the triangles (this was one of the subjects who performed the experiment three times). The numbers on the right give the slopes of regression lines for each of the four kinds of symbols (the shaded symbols were included with both the open and solid symbols, because the two kinds of simulations are obviously identical when the simulated velocity of ego-motion is zero).

If the subject had maintained the velocity relative to himself, the slope would be zero for both simulations. If he had maintained the retinal velocity relative to that of the most distant structure, the slope would always be -1. If he had maintained the local relative velocity it would have been 1 for the simulation of combined rotation and translation, and -1 for the simulated rotation. The average of the slopes for the two transitions (from faster to same perceived velocity [downward pointing triangles] and from same to slower perceived velocity [upward pointing triangles]) was determined for each subject and each kind of simulation (ego-rotation; combined ego-rotation and translation). These averages are shown in Fig. 6.

The three open symbols in Fig. 6 indicate what subjects would set if they relied exclusively on the target's velocity relative to themselves (absolute), relative to the adjacent surrounding (local) or relative to the most distant structure (horizon). It is evident that subjects do not rely exclusively on any one of these sources of information. Moreover, there are considerable differences in the extent to which subjects rely on retinal information. Our main interest was in the retinal component. If subjects combined either of the proposed sources of retinal information in a fixed manner with extra-retinal information, their data should fall on one of the dotted lines.

There is a clear tendency to rely on the most distant structure (negative values of the slope for simulated rotation and translation) for the retinal contribution, but the slopes for the combined rotation and translation are generally smaller (less negative) than those for simulated rotation. Two subjects show similar slopes for both conditions, one shows a slightly higher and three a slightly lower slope for the combined rotation and translation. The remaining four subjects show almost no influence of the background motion in the combined rotation and translation condition, although they were influenced by the background for simulated rotation. Again, subjects never experienced vection.

Discussion

The results of the second experiment confirm that motion of the images of distant structures provide the most important visual contribution when accounting for one's own motion (negative vertical values in Fig. 6). It is also evident that subjects do not rely exclusively on this measure.

When the whole background shifted at a single velocity to the left (simulating ego-rotation to the right), the perceived velocity of the target was increased by about half of the velocity of the moving background. This is the approximate magnitude of the influence of a moving background when subjects are asked to match velocities presented during separate intervals (e.g. Smeets & Brenner, 1995). In our previous studies in which subjects were asked to make judgements on changes in target velocity at the onset of simulated ego-motion (Brenner, 1991; Brenner & van den Berg, 1994), the influence of the moving background was considerably larger.

The apparently larger influence in the previous studies, and in the first experiment of the present study, is partly due to a shift in emphasis. Until now, we have emphasised that the range of velocities for which the target appeared to continue to move at the same velocity included the values at which relative velocity was maintained. However, this range usually extended asymmetrically around the value predicted from relative motion, with most of the range lying in the direction of the actual velocity (as in Fig. 2 of the present study). As a result, the slope of target velocity as a function of background velocity is less steep. The fact that the range of velocities for which the target was reported to appear to continue to move at the same velocity in this experiment (e.g. Fig. 5) often did not include the value one would expect on the basis of relative motion, is probably partly due to our explicit instructions to keep this range as small as possible.

The larger contribution of extra-retinal information in the present experiment may also have to do with the higher target velocity (27 deg/sec), although target velocity appeared to make little difference at lower velocities (6–12 deg/sec; Brenner & van den Berg, 1994). Alternatively, the differences may not really be due to extra-retinal information at all. The large projection screen has the disadvantage that it is impossible to keep the room dark enough to prevent subjects from seeing any stationary contours (such as the edges of the screen and texture on the floor in front of the screen). Such static contours should be irrelevant (assuming that we base our judgements of object velocity on motion of the structures that are perceived to be most distant), because these contours are always very close to the subject. However, several subjects explicitly reported that the visible border of the screen influenced their judgements. Presumably they were influenced to some extent by the target's final position on the screen.

A more important issue for our attempt to determine how retinal information is used to account for our own motion is why we often found a larger influence of background motion for simulated rotation alone, than for the combined rotation and translation. We propose several possible explanations.

First, the conflict between retinal and extra-retinal information is smaller for the simulated rotation. The simulated rotations in the combined rotation and translation are twice as large as those for simulated rotations alone, and they are accompanied by fast simulated translations (Table 1).

Second, the triangles are distributed at random on the floor. Whenever the most distant triangle is nearer than 600 cm, the influence of the simulated combined rotation and translation is reduced [because structures nearer than 600 cm move more slowly to the left; see Fig. 3(f)], but that of the simulated rotation is not [see Fig. 4(f)]. Moreover, subjects may misjudge which triangle is most distant, or use the average velocity of several distant triangles, which would decrease the influence of the simulated combined rotation and translation for the same reason.

Third, we may not ignore local motion altogether. In particular, a perceptual conflict may arise when the target stops moving relative to the local surrounding texture, and thereby becomes part of the static environment. For simulated rotation at 27 deg/sec and translation at -1.5 m/sec (see Table 1), the target would be expected to move to the right at 15 deg/sec, while the floor is also moving at 15 deg/sec to the right (see Fig. 5).

Taking these arguments into consideration, we conclude that six of our subjects' results are consistent with the hypothesis that motion is primarily judged by combining extra-retinal signals with the motion of the target's retinal image relative to that of the most distant structures. The other four subjects' results are less conclusive because they showed very little influence of background motion for simulated rotation and translation (although they showed a similar influence to that of the others for simulated rotation). We therefore conclude that whenever the visual surrounding does have an effect, this effect is dominated by the most distant structures. Undoubtedly, the extent to which subjects rely on visual information—when in conflict with extra-retinal information—depends on many factors (Brenner & van den Berg, 1994; Brenner *et al.*, 1996), so that differences between subjects are to be expected. The balance between retinal and extra-retinal information probably also depends on the target's velocity and the kind and speed of simulated ego-motion. We therefore do not wish to conclude anything about this balance, but only that the balance is primarily between extra-retinal information and retinal motion relative to that of distant structures.

GENERAL DISCUSSION

Retinal and extra-retinal information are combined when judging a moving object's velocity. We now show that the retinal contribution is dominated by motion of the target's retinal image relative to that of the most distant structure. This provides an estimate of the target's motion relative to ourselves, without requiring detailed knowledge of the target's distance or of our own motion.

Motion of the most distant structure's retinal image is used to estimate eye-rotation. This estimate is probably combined with an extra-retinal estimate of eye-rotation (Brenner & van den Berg, 1994). There is no evident need to estimate eye-translation. This may explain why eye-translation is not adequately accounted for on the basis of extra-retinal information: In the dark, during self-induced lateral ego-motion similar to the motion that was simulated in the present study, subjects systematically misjudged the lateral motion of a single light source (Gogel, 1982). The misjudgements of the targets' distances that would account for the errors were very different from the distances that subjects indicated when asked to point at the targets, so it is unlikely that the errors are (only) due to misjudgements of distance.

Despite the fact that we did not instruct the subjects on which frame of reference they should use, all subjects appeared to be judging object motion relative to the moving eye. This is in accordance with a similar experiment on perceived motion in depth during simulated ego-motion in depth (Brenner & van den Berg, 1993). In that experiment, subjects spontaneously judged the target's velocity relative to the eye (on the basis of the rate of expansion of the target and the vergence required to maintain fixation), completely ignoring the expansion of and changing disparity in the background (simulating ego-motion in depth). In fact, when we showed subjects their performance (in that study), and repeated the experiment with the explicit instruction to report on motion relative to the surrounding, they had difficulties with the task and performed very poorly.

Our results are consistent with some recent findings concerning the use of visual information to determine the direction in which we are heading. Although the extent to which we can determine our simulated direction of heading from visual displays in which combinations of eye-rotation and eye-translation are simulated is still a matter of some controversy (Warren & Hannon, 1988; Van den Berg, 1992; Royden *et al.*, 1992, 1994), there are clearly some conditions in which we can do so. This requires an ability to separate the retinal flow field into influences of translation and of rotation. One way to do so would be to consider all motion relative to (the motion of) the most distant structure, as here suggested for perceived object motion. This would at least partly account for the influence of eye-rotation, because the retinal motion of the most distant structures is least influenced by eye-translation. It will, however, give rise to considerable systematic errors if the most distant structure is nearby. Moreover, it requires independent information on structures' distances.

Several aspects of our ability to determine our direction of heading from the retinal flow field support this hypothesised mechanism of isolating the translational flow field. We can tolerate larger disturbances to the flow field when perspective (Van den Berg, 1992) or stereopsis (Van den Berg & Brenner, 1994b) provide information on structures' distances, than when they do not. Moreover, for simulated motion across a ground plane, limiting the visible range makes us misjudge our direction of heading in the way that is predicted by the use of retinal motion relative to that of structures at the horizon (Van den Berg & Brenner, 1994a).

The present results are also consistent with reports that the most distant structure (or the one that appears to be most distant) determines whether subjects experience circular vection when two structures of a display are in conflict; one moving at a constant angular velocity across the subject's field of view, and the other static (Brandt *et al.*, 1975; Ohmi *et al.*, 1987). Thus, estimates of eye-rotation based on the most distant structure appear to account for one's circular vection, as well as providing the basis for dealing with one's rotations when judging object motion and one's direction of heading. It provides an estimate of the object's *angular* velocity relative to ourselves, without requiring metric information on depth or on our own motion. This estimate will normally conform with estimates based on extra-retinal information. Obviously, in order to obtain an estimate of the actual velocity relative to ourselves, this measure of angular velocity must be combined with information on distance (Brenner, 1993; see Sedgwick, 1986 for a review on distance cues) and motion in depth (Brenner & van den Berg, 1994; see Regan *et al.*, 1986 for a review of cues for perceiving motion in depth).

REFERENCES

Brandt, T., Wist, E. R. & Dichgans, J. (1975). Foreground and background in dynamic spatial orientation. *Perception and Psychophysics, 17*, 497–503.

Brenner, E. (1991). Judging object motion during smooth pursuit eye movements: The role of optic flow. *Vision Research, 31*, 1893–1902.

Brenner, E. (1993). Judging an object's velocity when its distance changes due to ego-motion. *Vision Research, 33*, 487–504.

Brenner, E. & van den Berg, A. V. (1993). Perceived object motion in depth during simulated ego-motion. *Perception, 23 Suppl.*, 51–52.

Brenner, E. & van den Berg, A. V. (1994). Judging object velocity during smooth pursuit eye movements. *Experimental Brain Research, 99*, 316–324.

Brenner, E., van den Berg, A. V. & van Damme, W. (1996). Perceived motion in depth. *Vision Research, 36*, 699–706.

Cornilleau-Pérès, V. & Droulez, J. (1994). The visual perception of three-dimensional shape from self-motion and object motion. *Vision Research, 34*, 2331–2336.

Gibson, J. J. (1979). *The ecological approach to visual perception*. Boston, MA: Houghton Mifflin.

Gogel, W. C. (1982). Analysis of the perception of motion concomitant with a lateral motion of the head. *Perception and Psychophysics, 32*, 241–250.

Koenderink, J. J. (1986). Optic flow. *Vision Research, 26*, 161–180.

Land, M.F. (1992). Predictable eye–head coordination during driving. *Nature, 359*, 318–320.

Mergner, T., Rottler, G., Kimmig, H. & Becker, W. (1992). Role of vestibular and neck inputs for the perception of object motion in space. *Experimental Brain Research, 89*, 655–668.

Ohmi, M., Howard, I. P. & Landolt, J. P. (1987). Circular vection as a function of foreground–background relationships. *Perception, 16*, 17–22.

Post, R. B. & Lott, L. A. (1990). Relationship of induced motion and apparent straight-ahead shifts to optokinetic stimulus velocity. *Perception and Psychophysics, 48*, 401–406.

Raymond, J. E., Shapiro, K. L. & Rose, D. J. (1984). Optokinetic backgrounds affect perceived velocity during ocular tracking. *Perception and Psychophysics, 36*, 221–224.

Regan, D. M., Kaufman, L. & Lincoln, J. (1986). Motion in depth and visual acceleration. In Boff, K. R., Kaufman, L. & Thomas, J. P. (Eds), *Handbook of perception and human performance: Volume 1. Sensory processes and perception* (pp. 19.1-19.46). New York: Wiley Interscience.

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

Royden, C. S., Banks, M. S. & Crowell, J. A. (1992). The perception of heading during eye movements. *Nature, 360*, 583–585.

Royden, C. S., Crowell, J. A. & Banks, M. S. (1994). Estimating heading during eye movements. *Vision Research, 34*, 3197–3214.

Schwarz, U., Busettini, C. & Miles, F. A. (1989). Ocular responses to linear motion are inversely proportional to viewing distance. *Science, 245*, 1394–1396.

Sedgwick, H. A. (1986) Space perception. In Boff, K. R., Kaufman, L. & Thomas, J. P. (Eds), *Handbook of perception and human performance: Volume 1. Sensory processes and perception* (pp. 21.1–21.57). New York: John Wiley.

Smeets, J. B. J. & Brenner, E. (1995). Perception and action based on the same visual information: Distinction between position and velocity. *Journal of Experimental Psychology: Human Perception and Performance, 21*, 19–31.

Van den Berg, A. V. (1992). Robustness of perception of heading from optic flow. *Vision Research, 32*, 1285–1296.

Van den Berg, A. V. & Brenner, E. (1994). Humans combine the optic flow with static depth cues for robust perception of heading. *Vision Research, 34*, 2153–2167.

Van den Berg, A. V. & Brenner, E. (1994). Why two eyes are better than one for judgements of heading. *Nature, 371*, 700–702.

Warren, W. H. & Hannon, D. J. (1988). Direction of self-motion is perceived from optical flow. *Nature, 336*, 162–163.