JUDGING OBJECT MOTION DURING SMOOTH PURSUIT EYE MOVEMENTS: THE ROLE OF OPTIC FLOW

ELI BRENNER

Neuro-ethology Group, University of Utrecht, Limalaan 30, 3584 CL Utrecht, The Netherlands

(Received 27 August 1990; in revised form 4 March 1991)

Abstract We tend to follow moving objects with our eyes. To estimate their velocities, therefore, we must take account of our eye movements. During smooth pursuit, velocity judgements can be led astray by moving the background. Do we misjudge an object's velocity when the background moves because the additional shift of the background's image on the retina is interpreted as the result of additional motion of the observer rather than as motion of the background? In the present experiment, the traditional configuration of target and background was supplemented with a "floor of tiles" drawn in perspective directly under the "background". The motion of this new simulated plane was used to specify whether the additional retinal shift represents actual motion in the background, rotation of the observer's eyes, or observer locomotion parallel to the target. Moving the background clearly influenced the perceived velocity of the target. However, "specifying" whether the observer or the background had moved did not affect the outcome. For observer locomotion parallel to the target, the change in target velocity that is predicted by the optic flow depends on the perceived distance of the target. Nevertheless, presenting the target at different distances (by presenting different images to the two eyes) did not affect the subjects' settings. The results show that our judgement of objects' velocities does not depend on an assessment of our own movements on the basis of a global analysis of the optic flow.

Motion perception Optic flow Eye movements Psychophysics Smooth pursuit Velocity Spatial vision

INTRODUCTION

We often keep a moving object's image on the part of our retina with the highest spatial resolution by following it with our eyes. Except for allowing us to detect small detail, that may be essential for recognizing the object, this reduces the retinal blur that would occur if the image were to shift too rapidly across the retina. Determining the object's velocity, however, becomes more complicated if we move our eyes.

When we follow a flying bird with our eyes, its image hardly moves on the retina. The image of the trees behind the bird, however, moves in the opposite direction, with the angular velocity that is needed to follow the bird's motion. Nevertheless, we see the bird move and not the trees. Evidently, we take account of the fact that our eyes are moving.

Von Holst and Mittelstaedt (1950) proposed that the perceived object motion is ridden of effects of eye movements by "subtracting" the motion of the image that one would expect on the basis of (a copy of) the signals driving the eye muscles from the actual motion on the retina. Similar signals, could be used to account for other movements of the observer (Biguer, Donaldson, Hein & Jeannerod, 1988; Howard, 1986; Mack, 1986; Probst, Brandt & Degner, 1986; Wallach, 1985; Wertheim, 1990).

Velocity judgements can be led astray by moving the background. This has been explained by assuming that intentional eye movements are compensated for, as proposed by Von Holst and Mittelstaedt, whereas reflex-driven eye movements are not (e.g. Raymond, Shapiro & Rose, 1984). However, a hypothesis based on reflexes cannot explain why the perceived motion depends on which part of the display is seen as the background (Brandt, Wist & Dichgans, 1975; Ohmi, Howard & Landolt, 1987; Ohmi & Howard, 1988).

In the present study, I will examine another hypothesis. When an observer moves in a static environment, the image on his retina changes systematically. These systematic changes could be used to determine the displacements and rotations that his retina makes in relation to the environment. Such information could be combined with the motion of the target on the retina for judging the object's velocity relative to the static environment. The main argument against
this hypothesis, is that we perform quite well under experimental conditions in which the optic flow does not provide the necessary information; e.g. when judging the velocity of a moving light source in an otherwise dark room. However, in everyday life, we may mainly base velocity judgements on visual information.

When the extra-ocular muscles are paralysed, the perceived position of a target—in an otherwise dark environment—changes when subjects attempt to move their eyes. In a normally illuminated room it does not (reviewed in Matin, 1986). Apparently visual information suppresses extraretinal signals if a conflict arises. This is also the case under many other conditions in which subjects are confronted with conflicting visual and non-visual information on a target's position or velocity (Mack, 1986; Matin, 1986).

Lishman and Lee (1973) demonstrated that optic flow dominates our sensation of locomotion when visual information is in conflict with that of other senses. Visual information also dominates as an indicator of the observer’s displacement when assessing relative depth from motion parallax (Braunstein & Tittle, 1988; Rogers, Ono & Rogers, 1988). Warren and Hannon (1988) have shown that we can determine the direction in which we are heading from the optic flow alone. Moreover, this ability is independent of extra-retinal information on eye movements; instructing the subjects to maintain fixation on one (moving) point in the display, or providing the optic flow that would arise if they were to do so without the subject actually making any eye movements, had little effect on the accuracy with which the subject could indicate the direction in which he was heading.

Basing judgements on optic flow has the advantage that it is independent of the many movements that the observer makes that together form the final translation and rotation of the environment in relation to the observer’s eyes. However, in simple experiments with a single moving background [e.g. Fig. 1(b)], the observer cannot tell from the visual input alone whether the background has moved, or whether he himself has moved. This problem underlies the perception of ego-rotation in a rotating drum. In the present study, I examine whether modifying the image so that the optic flow suggests that the background has actually moved eliminates the influence of the background; whereas modifying it in a way that suggests that the observer himself has turned, or has moved parallel to the moving target, increases its effect.

**METHODS**

The experiments were conducted using an ATARI Mega ST 4 computer with an ATARI SM 125 white monochrome screen (71 Hz; 640 x 400 pixels); unless stated otherwise. Subjects looked at a 22 x 13 cm image with one eye from a distance of 35 cm. They had to look through a 15 mm dia hole in a “box”. Opposite this hole, the open side of the box fit tightly to the screen. The inside of the box was painted matt black. The room in which the tests were conducted was kept dark, to ensure that subjects had no additional visual points of reference. The stimulus was a 1x1 cm random pixel array (50% light; 50% dark) that moved from left to

---

Fig. 1 (facing page). Schematic representation of the stimuli used in the first experiment. The target (square with dots) moved from left to right across the screen (as indicated by the solid arrows at the bottom of the figure). After 500–700 msec of motion at a constant speed, the target gradually changed velocity (this took 100 msec), and continued at the new speed for 300–400 msec. The change in velocity was indicated by a tone. At the same time, the background too could start moving (open arrows). The subject was instructed to follow the target with smooth pursuit eye movements and to indicate whether the target moved faster, more slowly, or at the same speed after the tone. The target consisted of a random pixel array shifting across the similarly textured rectangle. A “floor of tiles” presented in perspective directly below the “wall” defined whether the additional motion of the background should be interpreted as motion of the “wall” (or of texture on the wall), or as the result of either translation or rotation of the observer. Six situations are shown: a target moving in the dark (a); a target moving across a frame within which the pixels move to the left (b; note that pixels disappear at the left, while new pixels appear on the right); the same stimulus with a stationary floor of tiles (c; simulating background motion to the left with no movement of the observer that cannot be accounted for by her eye movements); a target moving to the right while a similarly textured rectangle behind it moves to the left (d); the same stimulus with a floor of tiles shifting together with the textured rectangle (e; simulating additional rotation to the right on the part of the observer); and the same stimulus with the tiles moving in a manner simulating additional motion to the right on the part of the observer (f; note the deformation of the tiles).
right across the screen [Fig. 1(a)]. The luminance of the pixels was kept quite low: light pixels provided 3.0 cd/m² and dark pixels 0.015 cd/m²; higher levels proved to be unpleasant. The low level ensured that no other contours were visible.
The target appeared on the left half of the screen and started moving to the right. A tone warned subjects of the onset of motion. The subjects had been instructed to follow the moving target with their eyes. After between 500 and 700 msec, when the target had reached the centre of the screen, it gradually changed its speed (within 100 msec), and then continued at the new velocity for another 300–400 msec (human observers are known to be able to detect differences in velocity of less than 5% for targets shown for as short as 200 msec; McKee, 1981). The change in velocity was indicated by a second tone. Subjects had to report whether the target moved faster, slower, or at the same speed after the second tone. They were explicitly requested to ignore motion in the background. The target's velocity after the second tone depended on the subject's choices on previous presentations. A special staircase procedure was used to find two velocity settings for each experimental condition: the speed at which the target appears to accelerate and that at which a reduction in speed is observed. The different experimental conditions and the staircase procedure will be described in the two following paragraphs.

I examined the effects of various backgrounds and their movements on perceived velocity. Before the second tone, there were three possible configurations: a dark background [Fig. 1(a)]; a random pixel array [Fig. 1(b) and (d)]; or a similar random pixel array—that appeared to form the far wall of a perspective view of a room—with a simulated floor consisting of 3 rows of 8 black and white tiles directly beneath it [Fig. 1(c), (e) and (f)]. The wall filled the upper ⅔ of the image, and the floor filled the lower ⅓. After the second tone, both the random pixel array constituting the target's direct background (the "wall") and the simulated "floor" could move. To simulate actual background motion, the pixels within the wall shifted, whereas the floor remained still [Fig. 1(c)]. To simulate rotation of the observer, both the wall and the floor shifted across the screen [Fig. 1(e); the background moves as a whole, without changing]. Simulating a displacement of the observer gave a similar shift on the screen, except that the floor of tiles underwent a deformation that is related to the fact that displacement of the observer makes "nearby" contours shift more rapidly than "distant" ones. As controls I used the same displays without the floor of tiles [Fig. 1(b) and (d)]. Pixel motion within a stationary outline was used as a control for the simulated motion of the wall itself [Fig. 1(b) and (c)]. Motion of the whole random pixel array ("the wall") was used as a control for simulated motion of the observer [Fig. 1(d), (e) and (f)]. In the former case, pixels disappear at one side, and new pixels appear at the other. In the latter, the whole image simply shifts. In the vicinity of the target, the two conditions [Fig. 1(b) and (d)] are identical.

For finding the transition from no perceived change in velocity to an increase in velocity, the staircase procedure was as follows: if the subject reported that the target accelerated, target speed after the second tone was reduced by the current step size. If she reported that it either did not change its speed or that it moved more slowly after the second tone, target speed was increased by the same step size. The step size was then decreased to 80% of its former value. The initial step size was half of the target speed before the second tone (the reference speed). After the first step it was 0.8 times this speed; after the next step 0.64 times; and so on, until the step size reached a level that was negligible on our screen (0.08 pixels per frame or about 1 pixel every 180 msec). The 80% reduction in step size is a compromise between a high rate of convergence (maximal for a 50% reduction in step size) and repeating measurements to average perceptual variability. Each decision can be cancelled by an opposite decision on the following two steps. The setting onto which each staircase converges was taken as the transition point. The transition from no change to a decrease in velocity was determined in the same manner, except that reports of either no change in speed or of an increased velocity resulted in a lower velocity in the next presentation, whereas only reports of a decrease in velocity resulted in a corresponding increase. Presentations for all of the separate staircases (two staircases per condition) were intermingled so that the subject had no idea of which stimulus would follow.

Two precautions were taken to prevent subjects from estimating time or displacement rather than velocity. The first was the use of random presentation times. Both before and after the second tone, the presentation time was chosen at random, within the limits mentioned above. The second precaution was the use of random pixel arrays, so that the target could not be located after it stopped moving. In the initial experiment, the background moved at 0, ⅓ or 1 pixel per frame horizontally in either direction.
for a target reference velocity of \( \frac{1}{2} \) pixel per frame (about 6°/sec), and at \( \frac{1}{2} \) a pixel per frame in either direction for target velocities of \( \frac{3}{4} \), 2, and 3 pixels per frame (about 3, 9 and 12°/sec). A velocity of \( \frac{1}{2} \) a pixel per frame is actually a shift of 1 pixel every 2 frames; etc.

In a second set of experiments, similar stimuli were projected onto a large screen using a (Barco) video projection system (60 Hz; 640 x 200 pixels). Subjects looked at the screen with both eyes from a distance of slightly under 1 m (maximal luminance: 1 lx). The image filled 100° of visual angle horizontally and 89° vertically, rather than 35 and 22° respectively in the first experiments. Moreover, the random pixel array only filled \( \frac{1}{2} \) of the image vertically \( (\frac{1}{2} \) horizontally), whereas a floor and additional ceiling of tiles filled the rest. The initial target velocity was \( \frac{1}{2} \) pixel per frame (about 9°/sec at the centre of the screen). Background velocity was \( \frac{1}{2} \) and \( \frac{1}{2} \) a pixel per frame. The rest of the procedure was identical to that in the first experiment. The advantages of a large screen for simulating optic flow are obvious. A disadvantage is that the precision with which each setting is made is reduced, because pixel size limits precision, and a larger display results in larger pixels.

In a third experiment, the simulation of translation of the observer was repeated as in the initial experiment [Fig. 1(f)], but the depth at which the target moved was defined by stereopsis. To do so, images were presented to the left and right eyes in succession with LCD shutter spectacles (Neucom Electronic GmbH). The position of the target in each image depended on the simulated depth of the target (relative to that of the wall), the eye for which the image was intended, and the distance between the subject's eyes. The tiles were also presented to the left and right eyes with the appropriate disparities, confirming that they formed a "floor". The target was presented at 60, 70, 80, 90 and 100% of the distance to the wall, and moved at \( \frac{1}{2} \) pixel per frame. Background motion was always 1 pixel per frame in the opposite direction than the target. The target was shown (in depth) for 3 sec before starting to move, in order to give the subjects time to adjust to the disparity. The intensity of light pixels on the screen was increased to 50 cd/m², to compensate for the reduction in intensity caused by the spectacles.

**RESULTS**

Average results for all conditions and an initial target speed of \( \frac{1}{2} \) pixel per frame (6°/sec) are summarized in Fig. 2. The velocity at which the target appears to increase (open symbols) or decrease (solid symbols) its speed—in comparison with the velocity before the second tone—is shown as a function of the background velocity after the second tone. At velocities between corresponding open and solid symbols, the target does not appear to change its speed. The target's velocity before the second tone is indicated by the thick dashed line. The setting that would be required to maintain a constant relative velocity between the target and its direct surrounding is shown by the thick solid line. The shaded area depicts the range in which no change in velocity was reported when the background was dark [Fig. 1(a)].

![Fig. 2. Mean transitions between an increase and no change in perceived velocity (open symbols) and between no change and a decrease in velocity (solid symbols) for the 6 subjects that took part in this experiment. The thick dashed line indicates the target speed that subjects were required to match. The continuous line indicates the speed at which relative velocity remains constant. The shaded area shows the range of speeds for which subjects reported no change in velocity when the target moved in the dark [Fig. 1(a)]. Positive values on the horizontal axis indicate background motion in the same direction as the target. Negative values indicate motion in the opposite direction. The symbols represent motion within the background [diamonds; Fig. 1(b)], motion of the background [triangles pointing downwards; Fig. 1(d)], motion on the wall of a room facing the observer [triangles pointing upwards; Fig. 1(c)], simulated rotation of the observer in the room [squares; Fig. 1(e)], and simulated locomotion of the observer parallel to the target [circles; Fig. 1(f)]. The conditions represented by these symbols differ in the manner in which global optic flow predicts that the additional shifts in the background's image on the retina should be interpreted. The pursuit eye movement also shifts the image of the background, but this shift is accounted for by the subject's eye movements. One pixel per frame corresponds with about 4°/sec.](image-url)
Fig. 3. Transitions between an increase and no change in perceived velocity (open symbols) and between no change and a decrease in velocity (solid symbols) for the 6 subjects that took part in this experiment. Except for myself and one colleague (Bert), the subjects were volunteers who were not informed on the purpose of the experiment until after they were tested. I performed the whole experiment 3 times. Alexandra performed the experiment twice. All other subjects performed the experiment once. Symbols as in Fig. 2.

retinal displacements, rather than on an interaction between the target's displacement on the retina and extra-retinal information. However, the settings were very similar for all optic flow conditions (represented by the different symbols in Fig. 2). Contrary to the hypothesis proposed in the introduction, providing visual information specifying the origin of the moving background—in terms of either actual background motion or motion of the observer—does not change the influence that a moving background has on the perceived velocity. The only case in which there may be a modest effect was when the floor of tiles was stationary whereas the background moved in the same direction as the target [Fig. 1(c); triangles pointing upwards at positive values of background speed in Fig. 2]. The lack of systematic differences between conditions providing different global optic flow was evident for all six subjects.

Fig. 4. Data for various target speeds for the same subjects as in Fig. 2. The left part of the figure shows values for background motion in the opposite direction than the target at 0.5 pixels per frame (one pixel every 2 frames). The right side shows values for motion in the same direction as the target at that velocity. Symbols as in Fig. 2. One pixel per frame corresponds with about 4°/sec.
Object velocity during ocular pursuit

The large effect of background motion in all three experiments is evidence that the visual background plays an important role in judging object velocity. However, the perceived velocity appears to depend on the speed with which certain parts of the surrounding shift on the retina, rather than on the result of a complete analysis of the optic flow. Despite clear differences between the settings made by individual subjects, the settings made under the different optic flow conditions were very similar for each subject.

Actually, optic flow can only provide information on the relative motion of objects with respect to the observer. However, if the whole environment moves systematically with respect to the observer, the optic flow will be interpreted as movement of the observer himself (Lishman & Lee, 1973). In the present experiment, the subjects were not fooled by the optic flow; they never had the feeling that they themselves were moving, and usually noticed when the background moved. One could argue that the fact that subjects noticed that the "tiles" moved (for simulated additional ego-rotation) and did not have a sensation of ego-motion (for simulated displacement of the observer) implies that the stimulus was not adequate to test the hypothesis. However, this argument can easily be dismissed. The effect of the background on perceived target motion was quite evident, although interpreting the image (correctly) as actual background motion would predict no effect. Apparently, judgements of object motion are independent of the sensation of ones own motion.

Can the results be explained without dismissing global optic flow as a factor in object velocity judgements? The perceived target velocity appeared to depend on its motion relative to the "wall". This would be expected on the

**DISCUSSION**

**Fig. 5.** Data for stimuli presented on a large screen (see text). The figure shows the average of my data and that of five naive subjects. I performed the experiment 3 times, and the other subjects (volunteers who were not informed on the purpose of the experiment until after they were tested—one of whom had also taken part in the previous experiment) either performed the experiment once (2 subjects) or twice (3 subjects). Symbols as in Fig. 2. One pixel per frame corresponds with about 12°/sec.

**Fig. 6.** The effect of the distance to the target on its perceived velocity for simulated translation of the observer parallel to the target [Fig. 1(f)]. Relative distances were defined by stereopsis. The background moved at 1 pixel per frame in the opposite direction than the target. Symbols show the mean transitions between an increase and no change in perceived velocity (open circles) and between no change and a decrease in velocity (solid circles) for the 6 subjects that took part in this experiment (myself and 5 naive subjects). Each subject performed the experiment 3 times. The thick dashed line indicates the target velocity that subjects were required to match. The continuous line indicates the setting at which velocity relative to the direct surrounding remains constant. The curve shows the velocity that would be expected on the basis of the translation of the observer (suggested by the optic flow) and the depth of the target (suggested by stereopsis).
Fig. 7. The effect of moving either the “inner” (solid lines) or the “outer” (dotted lines) background on the perceived velocity for 5 of the subjects that took part in this experiment (the two lines indicate the upper and lower limits of the range of settings for which no change in velocity is perceived). The thin dashed and solid lines indicate the target velocity that subjects were required to match and the settings required to maintain the relative velocity respectively. Background speed always refers to the moving background (either the direct or the more distant surrounding). Each subject performed the experiment once. Six subjects were only (e.g. Simone) or mainly (e.g. Ingeborg) influenced by motion of the direct surrounding; three subjects were influenced to a similar extent by motion in either area (e.g. Peter-Jaap); three subjects were mainly or exclusively influenced by motion in the more distant surrounding (e.g. Jos); and one subject was hardly influenced at all (Arjan).

basis of optic flow for simulated rotation of the observer. It would also be expected for simulated translation parallel to the target motion, as long as the target is at the same distance as the wall. The closer the target is to the observer, the larger the expected effect (thick curve in Fig. 6). The only way to explain the absence of an effect of target distance in the third experiment, without dismissing global optic flow analysis as the basis for judging the target’s velocity, is by assuming that we do not use information from stereopsis to analyse the optic flow; despite recent evidence to the contrary (Roy & Wurtz, 1990). If so, however, the depth of the target is undefined. It is unclear, therefore, why the results should reflect the assumption that the target is at the same distance as the wall.

For simulated background motion, a similar situation arises. The results suggest that subjects consider the “wall” to be stationary. A consequence is that the floor must be moving. If so, we would expect its image to be distorted in the same way as it is for simulated ego-translation, because nearer contours should shift faster than more distant ones. This is not so. The velocity profile suggests that the whole floor is at the same distance from the observer; contradicting the information from perspective. In analogy to the case for stereopsis described in the preceding paragraph, we would have to assume that information from perspective is also not used when analysing the optic flow. Even so, however, there is still no reason why the wall should be considered stationary, rather than the floor (and ceiling). Moreover, I have some evidence that depth information derived from perspective is essential for the results.

It has been shown for both visually induced circular vection (Ohmi et al., 1987) and visually induced motion in depth (Ohmi & Howard, 1988) that the perceived background determines the sensation of ego-motion. Large, peripheral fields that appear to be far away are likely to be perceived as the background. Actual distance,
plane of focus, and whether the field is followed with tracking eye movements or not, are not important (Ohmi et al., 1987). Even the large, peripheral, stationary surfaces in the second experiment of the present study, hardly decreased the effect of the direct surrounding on perceived motion [triangles pointing upwards in Fig. 5, corresponding to the situation in Fig. 1(c)]. Apparently the wall was considered stationary, because it was seen as the most distant (as suggested by perspective).

In an additional test, I examined what would happen if I eliminated perspective from the display, making the depth relationships ambiguous. Two conditions were compared. In the first condition, the stimulus was identical to that shown in Fig. 1(b), except that the “wall” was surrounded by an additional stationary random pixel array. In the second condition, the “wall” was stationary, whereas the pixels in the additional surrounding image shifted. The procedure was as in the first experiment. Six subjects’ settings were still mainly or exclusively affected by motion in the adjacent part of the background (first condition). However, three subjects’ settings were mainly affected by motion in the more peripheral background (second condition); three subjects’ settings were affected similarly in both conditions; and one subject’s settings were hardly affected in either (Fig. 7). This variability in responding was not found in the previous experiments (see Fig. 3), presumably because perspective had indicated which surface was most distant. This demonstrates that the results of the other experiments were not simply due to perceived velocity depending on local relative speed; the more peripheral background affected the settings of about half of the subjects tested in this additional experiment (without perspective).

Perspective, optic flow, and stereopsis all support seeing the “wall” as most distant. The fact that motion relative to this—most distant—surface explains the results in all cases, confirms that the most distant surface is presumed to be static, and implies that motion of this surface is always presumed to be due to ego-rotation (even when the global optic flow indicates that this is not the case). This does not support a role for global optic flow in judging object velocity.

Despite the fact that subjects were explicitly asked to estimate absolute motion, ignoring motion in the background, the perceived velocity usually depended on relative motion. For backgrounds moving in the opposite direction than the target, this appeared to be exclusively so. For background motion in the same direction as the target, maintaining the relative velocity between target and background did not always prevent subjects from perceiving changes in target velocity (open symbols below the thick solid line in Fig. 3). The reasons for not always relying on local relative velocity will be examined in a following paper.

Acknowledgements—I wish to thank my friends and colleagues who volunteered to take part in the experiments, Raymond Würt for help with the video projection system, and Wim van de Grind and Bert van den Berg for reading and suggesting improvements to the manuscript. This study was supported by the Dutch Foundation for Psychological Research (PSYCHON grant 560-262-043).

REFERENCES


