Retinal and extraretinal contributions to visual stationarity

Mark Wexler[1]
Ivan Lamouret

Laboratoire de Physiologie de la Perception et de l'Action
Collège de France, 11 pl. Marcelin Berthelot
75005 Paris, France

November 22, 2000

[1]Corresponding author. E-mail: wexler@ccr.jussieu.fr, Tel: +33.1.44.27.16.24, Fax: +33.1.44.27.13.82.

Abstract

There are two ways in which moving observers could judge whether three-dimensional objects are stationary in an earth-fixed—rather than observer-relative—reference frame. An observer may determine the relative motion of an object as well as his or her own motion relative to an earth-fixed reference frame, and then compare the two: the extraretinal criterion. Alternatively, the observer could heuristically assume that the background in the image is stationary, and decide the stationarity of a foreground object on purely retinal criteria. We performed two experiments to determine which of these criteria is actually applied to make stationarity judgments. In the first experiment we show that both retinal and extraretinal criteria for stationarity are in themselves sufficient, but that retinal criteria yield stationarity judgments that are accurate but imprecise, extraretinal criteria are precise but sometimes inaccurate, and that the combination of the two is both accurate and precise. In the second experiment, we show that when both types of stationarity criteria are available, both are utilized; the relative weights, however, have high inter-individual variations.

1 Introduction

Moving observers can usually distinguish those objects that are stationary from those that are moving—not with respect to themselves, but with respect to an earth-fixed reference frame. Although this distinction seems to be made effortlessly, it is computationally difficult, in that all retinal input is relative to eye position, rather than to any external reference frame. The perception of an object as stationary despite movement in the visual image due to one's own displacement is a form of spatial constancy. The perception of object stationarity—and, more generally, the perception of object motion with respect to an earth-fixed reference frame, in spite of one's own movement—has unquestionable ecological value. Loss of the perception of stationarity—such as occurs when wearing glasses or prisms prior to adaptation, or in neurological cases (Haarmeier et al., 1997)—may have dramatic effects.

We can categorize self-displacements into two types: eye rotation and eye translation. Optic flow induced by eye rotation is easier to deal with, at least computationally, since it causes a field-wide shift in the retinal array that depends only on the movement itself, and not the 3D structure of the scene. Visual stability in the presence of eye rotations has been studied for quite some time (Descartes, 1664; von Helmholtz, 1867); for a recent review, see (Wertheim, 1994). The case of eye translation has received less attention, and is also computationally harder, since the resulting retinal shift depends on both the observer's displacement and on the 3D layout of the environment ('motion parallax'). In this article we examine the criteria for the perception of stationarity in the presence of eye translation.

In recent work, we have shown that stationary objects play a special role in spatial vision (Wexler et al., 2000). We will define an object as (strictly) stationary if its every point has constant position in a reference frame that is fixed to the earth. A notion related to stationarity is rigidity; the assumption of object rigidity is assumed to play a key role in the extraction of structure from motion (Ullman, 1979; Todd, 1982; Koenderink, 1986): a rigid object is defined as one in which the distance between any two points does not change. Therefore, an object that is stationary is also rigid, but not all rigid objects are stationary. In the above-mentioned work, we showed that the rigidity and stationarity criteria are intimately intertwined: stationarity increases the weight of the rigidity assumption (i.e., of motion cues to depth) relative to other depth cues (Wexler et al., 2000).

The retinal stimulus that leads to the perception of a moving object in a non-moving observer, may, in some cases, lead to the perception of a *stationary* object in a *moving* observer. The determination of object stationarity therefore cannot be *directly* extracted from retinal input without either integrating extraretinal information, or making heuristic assumptions about the scene (Wallach, Stanton, and Becker, 1974).

1.1 Extraretinal criteria

The only general method to determine whether a rigid object is stationary with respect to an earth-fixed reference frame is to compute the motion of the object with respect to the eye, as well as the translation of the eye with respect to the earth, and to subtract one from the other. The extraretinal information that contributes to the determination of self-motion may include pre-motor and motor, vestibular, and other proprioceptive signals concerning eye and head movements (Crowell et al., 1998). Thus, in the most general case, extraretinal information must be used in the determination of stationarity, and we shall call this subtraction method the extraretinal criterion (ERC). Of course, the movement of the eye in space may be extracted from optic flow (for recent reviews, see (Hildreth and Royden, 1998; Warren, 1998)), but the algorithms necessarily *assume* stationarity of the environment, and therefore do not generalize to non-stationary environments (we will address this issue in the following section).

A two-dimensional version of the ERC is the proposed mechanism that extraretinal eye position information helps stabilize the visual world during eye rotations (Mittelstaedt, 1990; Wertheim, 1994). The existence of such a mechanism is supported by extensive psychophysical data: for instance, in a paradigm first devised by Mach, it has been shown that efforts to move the eyes by subjects with paralyzed eye muscles lead to perception of visual motion in darkness (Stevens et al., 1976). Moreover, there is also neurophysiological data that shows that the visual fields of certain parietal cells in macaques shift in anticipation of eye movements, by precisely the amount necessary to keep the visual world stable (Duhamel, Colby, and Goldberg, 1992). However, this two dimensional stabilization mechanism seems not to be very accurate, as demonstrated by the Filehne illusion (Filehne, 1922; Mack and Herman, 1973), in which, during ocular pursuit, a object that is actually stationary seems to move opposite to the eye movement. Interestingly, the Filehne illusion can be modified or even reversed by a visual background (Wertheim, 1987; Haarmeier and Thier, 1996).

For the ERC to be a plausible method in the determination of stationarity in actual living systems, it must be shown that extraretinal information can contribute to the perception of three-dimensional structure and motion. This has been demonstrated in a number of different contexts. Extraretinal self-motion information is incorporated in judgments of object motion and absolute distance in three dimensions (Gogel and Tietz, 1973). In the heading-from-optic flow paradigm, extraretinal information has been shown to lead to more accurate heading estimates (Royden, Crowell, and Banks, 1994; Crowell et al., 1998). In the extraction of 3D structure from optic flow, extraretinal information in active observers can disambiguate symmetries of first-order flow (Rogers and Rogers, 1992; Dijkstra et al., 1995; Wexler, Lamouret, and Droulez, 2000), and lead to better spatiotemporal integration (van Damme and van de Grind, 1996). Finally, it has recently been shown that extraretinal information can alter the relative weights of multiple depth cues, and in particular to increase the weight of motion cues—but only if the object specified by the motion cues is stationary (Wexler et al., 2000). Thus extraretinal information is indeed integrated in the perception of structure from motion, and can in principle be used to compensate for self-motion in the perception of stationarity. Wallach argued that the spurious motion that is sometimes seen when translating the head about an object whose 3D structure is misinterpreted—such as a realistic perspective drawing, or a concave mask seen as convex—is evidence for self-motion compensation from extraretinal sources (Wallach, Stanton, and Becker, 1974).

However, one does not have to look far to find phenomena that demonstrate that the visual system does not always perform the general ERC computation. In the autokinetic effect, for instance, an isolated, stationary point in an otherwise dark visual field is perceived by the stationary observer to undergo significant absolute displacement (Koffka, 1935). Thus, in this simple case, the ERC for stationarity breaks down in the absence of a visual context. In vection, a large, slowly moving object is first perceived veridically; then gradually, the object's motion 'spills over' into a perception of self-motion (Dichgans and Brandt, 1978). If vection saturates, a non-moving observer actually perceives the moving field as stationary, and himself as undergoing motion in the opposite direction. Given these phenomena, it is difficult to imagine that the extraretinal criterion could be the *sole* determinant of stationarity.

1.2 Retinal criteria

The misperception of moving objects as stationary by an observer experiencing vection suggests that in some cases the visual system may simply *assume* that the visual background is stationary. Such a heuristic assumption would greatly simplify the determination of stationarity. Extraretinal information would no longer be required. Instead, optic flow could be used to compute the relative motion of a foreground object of interest, as well as that of the background: if the two differ, the foreground object would be perceived as non-stationary. In a cluttered scene, the "most distant structure" (Brenner and van den Berg, 1996) could be used as a reference. We shall call this heuristic the retinal criterion (RC), to remind the reader that it needs no extraretinal input.

The RC is a generalization of a rule proposed by Duncker, namely that in a visual scene with an object and a surround, or a foreground and a background, the background is assumed to be stationary (Duncker, 1929). Duncker's rule is motivated by the induced motion illusion that bears his name, and subsumes various center-surround motion distortions. The RC for stationarity would amount to a 3D generalization of Duncker's rule: formally, an object would be perceived as stationary if and only if it and the background move as one rigid body. It has been shown that there exist simple criteria to discriminate rigidity from optic flow, but that actual perception or discrimination of rigid motion by human observers is not very precise (Hogervorst, Kappers, and Koenderink, 1997).

In the absence of visual background, extraretinal information determines stationarity when the eyes rotate in darkness (Stevens et al., 1976), as mentioned above. However, as soon as a visual background is present, an RC-like phenomenon called *visual capture* occurs: the motion of the foreground object relative to the background is seen as absolute movement, and extraretinal information seems to be ignored. Visual capture has been found in both curare-induced paralysis (Matin et al., 1982) and eye-press studies (Stark and Bridgeman, 1983). Another type of visual capture occurs when the ERC and RC provide contradictory information about eye rotation (Wallach, Bacon, and Schulman, 1978); when, say a vertically moving object is tracked by the eyes with a horizontally moving background, diagonal motion is perceived, in contradiction to the extraretinal signal (Wallach, Bacon, and Schulman, 1978). A visual background, when present, thus serves as a reference for two-dimensional stationarity, overriding extraretinal information.

The assumption of background stationarity also underlies algorithms that are hypothesized to compute self-motion from optic flow (Gibson, 1950; Hildreth and Royden, 1998; Warren, 1998). This assumption is reasonable for the visual system to make, i.e., most of the time it would result in veridical perception. The criterion would not be applicable in situations where only the foreground is visible (e.g., illuminated objects at night), or lead to wrong results when the background is not actually stationary (e.g., a thick snowstorm in high winds (Chatziastros, Cunningham, and Bülthoff, 2000)). Whether this assumption is actually made by the visual system is an empirical question, to be addressed below.

1.3 Our experiments

The problem of how the world appears stable despite two-dimensional motion accompanying eye rotations has been studied at least since Descartes. What has been learned is that extraretinal information can be used to maintain the stability of the visual world; in the presence of a background, however, the human visual system can also judge stationarity by comparison to the visual background, ignoring the extraretinal information. However, the objects making up the real world are not the lines and dots in the frontoparallel plane that have been used in the above-mentioned studies, but surfaces with various 3D orientations. Eye movements in space are not only the pure rotations that have been studied (with the subject in a head rest), but often include translations. In such a case, two dimensional motion is of little use as a stationarity criterion; what matters is the 3D orientation of surfaces.

Let us consider the case of stationary three-dimensional objects. When we move past such an object while fixating a point on its surface, the retinal image is that of a surface that turns about an axis passing through the fixation point. Since the object is stationary, that rotation exactly compensates (in its angle and axis) for our movement. In the following experiments, we examine the criteria for stationarity in three dimensions by varying the angular speed of rotation of the surface relative to the observer's motion. Hans Wallach carried out studies of stationarity judgement of points (Wallach and Kravitz, 1965) and surfaces (Wallach, Stanton, and Becker, 1974), but the question of retinal versus extraretinal criteria was not addressed in these studies. Recent studies of stationarity judgements have also been reported (Jenkin et al., 2000).

In order to use the ERC to determine stationarity, the observer must evaluate and compare two angles (or angular speeds): that of his or her own rotation about the stationary point, and that of the object's rotation. The self-rotation can be calculated from several different extraretinal sources. For instance, it is equal to the absolute rotation of the eyes in space needed to fixate the stimulus (i.e., the rotation of the eyes with respect to the head minus the rotation of the head in space). The object's orientation and rotation can be obtained from either of the two depth cues available, motion or perspective. In order to use the RC, on the other hand, the visual system has only to determine the rotation of the surface relative to the visual background.

If the surface is the only object visible (as will be the case in one of the conditions in Experiment 1), the ERC is the *only* way to determine stationarity. On the other hand, if a visual background is present, the

RC can be applied. The two experiments to be described are an attempt to determine which of these criteria is actually employed by people making stationarity judgments. In the first experiment, only the ERC can be used in one condition, only the RC in the second condition, and both criteria are available in the third condition. In the second experiment, both criteria are available, but are in conflict with each other.

2 General methods

2.1 Apparatus

Subjects' translational eye displacement was measured by a special-made precision mechanical head tracker, which features sub-millimeter precision and time lag inferior to the sampling frequency used (Panerai et al., 1999). Eye position was sampled on-line at the frequency of our display, 75 Hz (using a National Instruments PCI-6602 data acquisition card) by a PC computer (Pentium II 400 MHz, Dell), which also controlled the visual stimuli. Stimuli were displayed on a flat, semi-translucent screen using a BARCO 1209 video projector (placed on the opposite side of the screen from the subject), driven by a Matrox G400 video card at a resolution of 1600×1200 pixels (image size 243×189 cm, yielding a spatial resolution of about 6.5 arcmin/pixel at typical observer distance of 80 cm), and at a vertical refresh rate of 75 Hz. Screen geometry was calibrated so that it nowhere deviated from a physical reference grid by more than one pixel.

2.2 Stimuli

Stimuli were simulations of virtual objects situated in front of the observer, polar-projected on each monitor frame (13.33 msec) for the position of the subject's dominant eye (the other eye was blindfolded), as read by the head tracker on the previous frame. The virtual objects making up the stimuli will be described in the *stimulus reference frame*, whose origin is the point on the screen closest to the subject's eye on the first frame of the trial, whose $x-$ and $y-$axes are the horizontal and vertical directions on the screen (directed rightwards and upwards, respectively, from the subject's point of view), and whose $z-$axis is perpendicular to the screen, pointing at the subject. Unless otherwise stated, all lengths will be given in centimeters.

All stimuli contained a central planar grid, made of 10×10 square cells, each 5 cm wide. In its central position, the center of the grid was at the origin, its slant 45°, and its tilt either 0° or 180°.[1] The grid texture on the surface was oriented so that all lines were either vertical or horizontal. After projection, the grid was drawn on the screen as 1-pixel-thick green lines.

Some stimuli also contained a background reference. This was a 'virtual room', comprised, from the subject's point of view, of a floor, ceiling, left, right and rear walls, surrounding the central grid. The width, height and depth of the room were 160 cm, 160 cm and 60 cm, respectively. The rooms was a parallelipiped, with the floor and ceiling in the $y = -80$ and $y = +80$ planes, the left and right walls in the $x = -80$ and $x = +80$ planes, and the rear wall in the $z = -60$ plane. The five surfaces were textured with

[1]Slant (σ) and tilt (τ) are defined so that the normal of the surface is $(\sin\sigma\cos\tau, \sin\sigma\sin\tau, \cos\sigma)$.

square grids, 20 cm wide, with all lines parallel or perpendicular to the edges. After projection, the background reference was drawn on the screen as 1-pixel-thick red lines. Because the origin of the stimulus reference frame depended on initial subject eye position, its height with respect to the screen varied with subjects' height. For very tall or short subjects, small parts of the top or bottom of the background reference would have projected outside the screen, and were therefore not drawn.

Finally, a fixation point was placed at the origin, drawn as a white circle with radius of 1 pixel. When the background reference was present, the central grid occluded it at any intersections. The fixation point was drawn last, so that it occluded all other stimuli. Other than the stimuli, the screen background was black, and the experiment was performed in the dark.

2.3 Task

The subject's task was to determine the direction of rotation of the central stimulus. More precisely, the subject had to determine whether the edge of the stimulus surface closest to the subject rotated to the subject's left or right. The reference frame for the task (earth-fixed, or with respect to a stationary or possibly moving visual background) depended on the experiment and condition. The subject responded at the end of trial, after the stimulus was no longer visible, by inclining a joystick either to the left or to the right. Since the subject's own motion was always from right to left, the objectively correct response was "right" for gains $\gamma < 0$, and "left" for $\gamma > 0$.

2.4 Observer and stimulus motion in active trials

Observers performed the experiment while standing in front of the screen on which stimuli were projected from the back. A trial began when the observer's dominant eye was within 10 cm of the ideal starting position. The distance of the ideal starting position was 80 cm from the screen, and its height was adjusted for each subject's typical eye height while standing. On the first frame of the trial, the point on the screen directly opposite the subject's eye was calculated, and subsequently used as the origin of the stimulus reference frame for that trial (see above).

At the beginning of an active trial, only the fixation point was displayed, which the subjects were instructed to fixate throughout the trial. The subject then moved his or her upper body to the right. When the subject's position[2] along the x-axis reached 10 cm, a brief tone was sounded, which was the signal to reverse movement direction, i.e., to start moving to the right. As soon as the x-position again fell below 10 cm, the stimulus was displayed. The subject was instructed to continue moving left, horizontally and parallel to the screen, as smoothly as possible, until a second tone was sounded, which occurred when the x-position fell below -10 cm. At this point the stimulus disappeared, the subject moved to the right, back to the starting position, and gave the response using a joystick. When the joystick was released, the next trial began. The eye trajectory was recorded in a data file, for possible use in stimulus generation in a subsequent passive trial.

[2] In the reference frame used here to refer to the position of the optic center of the subject's dominant eye (which we shall refer to as the 'eye position'), the origin is the initial eye position; the x-axis is horizontal and parallel to the screen, the y-axis is vertical, and the z-axis perpendicular to the screen.

On active trials, the central stimulus (and possibly the visual background) were rotated about the same axis as the subject's rotation with respect to stimulus origin, with the rotation calculated as follows. Let $\mathbf{r_0}$ be the initial position of the dominant eye, and $\mathbf{r}$ be its position at a given moment during the same trial, both measured in the stimulus reference frame. The subject has thus performed a rotation about the (stationary) stimulus center, with the axis $\hat{\mathbf{A}}$ of the rotation parallel to $\mathbf{r_0} \times \mathbf{r}$ (and passing through the origin, i.e. the stimulus center), and the angle of rotation given by $\theta = \arccos(\mathbf{r_0} \cdot \mathbf{r}/\|\mathbf{r_0}\|\|\mathbf{r}\|)$. At that moment during the trial, those parts of the stimulus that were non-stationary were rotated about the same axis ($\hat{\mathbf{A}}$), by an angle $\gamma\theta$, that is the same angle as the subject's rotation multiplied by a gain, γ.

2.5 Passive stimuli

In each trial in passive conditions, subjects experienced the same rotational optic flow as in a previous active trial, but without performing head movement. The subject was instructed not to make any head movements during the trial. At the beginning of the trial, the position of the subject's dominant eye, $\mathbf{r_0}$, was measured. As in active trials, the stimulus center for the passive trial was the point on the screen closest to $\mathbf{r_0}$. On each frame of the trial, the position of the subject's eye *for the corresponding frame during the corresponding active trial*, $\bar{\mathbf{r}}$ was read from a data file (with $\bar{\mathbf{r}}_\mathbf{0}$ the initial eye position during the active trial). As in active trials, the rotation $\overline{\Omega}$ about the (active) stimulus center that transformed $\bar{\mathbf{r}}_\mathbf{0}$ to $\bar{\mathbf{r}}$ was calculated. Any object movement (such as rotation by angle $\gamma\overline{\theta}$) was applied to the stimulus, followed by the *inverse* rotation, $\overline{\Omega}^{-1}$. Finally, the stimulus was projected for the eye position $\mathbf{r_0}$. This procedure ensured that the subject experienced the same rotational optic flow in the passive trial as in the corresponding active trial, but without self-motion.[3]

2.6 Data analysis

The raw data for each subject consisted of "left" or "right" responses for different gains in different conditions. These were first converted into fractions of "left" responses, as a function of gain (γ), for each subject in each condition. The fractions were then fitted (minimizing RMS) to logistic functions, $1/(1+e^{-z})$, $z = (\gamma - \gamma_0)/w$, where γ_0 is the bias of the sigma function (the 'point of subjective equality', where frequencies of "left" and "right" responses are 0.5), and w a measure of the width of its transition. The extent to which bias is close to zero would reflect the accuracy of stationarity judgements, whereas low width would reflect their precision. The fits were performed on individual subject data; performing fits on data averaged over subjects would have confounded precision and inter-subject bias variation. Conditions in which fitted bias fell outside of the range of experimental values for γ were deemed unreliable and were discarded from further analysis. The remaining fit parameters (biases and widths) were averaged over subjects to determine mean values in each condition; inter-subject variation was used in statistical tests. The threshold for statistical significance was taken

[3]The optic flow in active trials also had a small expansion/contraction component, due to small changes in eye-stimulus center distance during the subject's movement. These tiny expansions/contractions were not informative for the task, and were omitted from the optic flow in passive trials.

as 0.05.

3 Experiment 1

In the first experiment, we controlled the sources of information that are required to apply the extraretinal and retinal criteria for stationarity. In order to do so, we used a stimulus in which both channels of information were present, and two other stimuli, in each of which one of the two channels was removed. In the active frame condition (ACT-FRAME), observers performed head motions about a virtual stimulus surface that underwent a small rotation (relative to an earth-fixed reference frame) in synchrony with the subject's movement (see Fig. 1). In this condition, the stimulus was accompanied by a visual reference background (a 'virtual room') that remained stationary in space, regardless of the motion of the stimulus surface. The observer's task was a two-alternative forced choice judgement of the direction of stimulus rotation; performance on this task—more precisely, the width of the transition—was a measure of the accuracy of perception of stationarity. In the ACT-FRAME condition, therefore, the observer could use both retinal and extraretinal information in judging stimulus stationarity.

The active-object condition (ACT-OBJ) was identical to ACT-FRAME, save for the absence of the reference background. Thus, retinal information was no longer available, and only extraretinal information could be used in judging stationarity. Conditions ACT-OBJ and ACT-FRAME are similar to experiments performed by Wallach and his students (Wallach, Stanton, and Becker, 1974), who used a movement/stationarity discrimination task

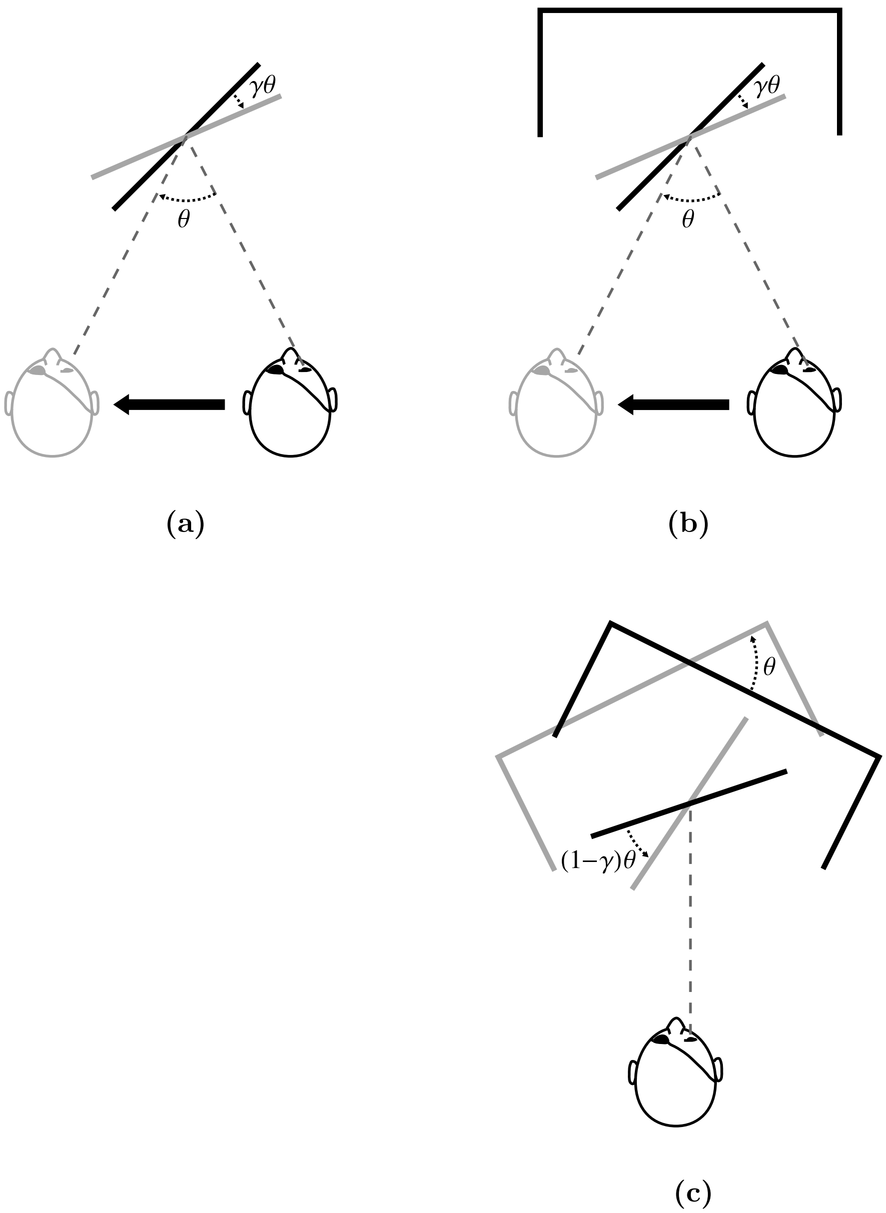

Figure 1: Simulated geometry and subject movement in Experiment 1. **(a)** In the ACT-OBJ condition, the observer moves from right to left while a surface object rotates about the same axis by the same angle multiplied by a gain, γ. An earlier stimulus, for the head position on the right, is shown in black, while a later stimulus, for the left head position, is in gray. **(b)** The ACT-FRAME condition is similar to ACT-OBJ, except for a visual frame that remains stationary. **(c)** The optic flow in the PASS-FRAME condition is the same as in ACT-FRAME, but the observer does not move.

(we shall compare our results with theirs below). In the passive condition (PASS-FRAME), on the other hand, the extraretinal information was removed by having a non-moving observer experience the same optic flow as in ACT-FRAME, with the task of judging the relative motion between stimulus and frame, rather than the absolute motion of the stimulus. In terms of optic flow, therefore, both stimulus and task are identical in ACT- and PASS-FRAME conditions.

If only retinal information is used in stationarity judgements, we predicted that performance in the ACT-FRAME condition, where such information is available, would be much better than in ACT-OBJ, where it is not. Moreover, the performance in PASS-FRAME, where the same retinal information was available, would be as good as in ACT-FRAME. On the other hand, if only extraretinal information is used to judge stationarity, performance in ACT-OBJ would be about the same as in ACT-FRAME, since the same extraretinal information is available in the two cases. Performance in PASS-FRAME, though, would be much worse, since the extraretinal information was no longer available in that condition.

We note in passing the reason we chose our particular task, rotation direction discrimination. Asking the subject to perform a stationary/moving discrimination would have been more direct, but results on such a task could depend on subjective limits for what is considered as "stationary", limits that could vary from subject to subject, or even from trial to trial. The widths of the resulting psychophysical curves would be difficult to interpret, leaving us only biases as a truly meaningful measure. Instead, we chose to have subjects discriminate the direction of rotation, a task that is free of such subjective elements. We could therefore interpret the width of the resulting psychophysical curve as the precision of stationarity detection.

3.1 Methods

The techniques used are described in the General Methods section. The experiment consisted of three main blocks, each corresponding to one of the three conditions. Subjects were randomly assigned to two groups, which differed in order of conditions. In Group 1, the order of the three blocks was ACT-OBJ, ACT-FRAME, PASS-FRAME; in Group 2, the order was ACT-FRAME, ACT-OBJ, PASS-FRAME.

The slant of the surface was always 45°, and two tilts, 0° and 180°, were used.[4] The visuomotor gain γ took one of six values: -0.5, -0.3, -0.1, 0.1, 0.3, 0.5. The design was factorial, with each condition condition repeated 10 times, which yielded 120 trials in each block. Preceding each block, subjects were given a practice block of 20 trials in the given condition, to familiarize them with the stimulus and procedure. A gain of 0 was used in the practice blocks.

Ten subjects (7 men and 3 women, ages $19 - 33$ years, mean age 27), volunteered to participate in the experiment, including the two authors and eight others who were naive to the experimental goals and hypotheses.

3.2 Results

Raw data, averaged over all subjects, are shown in Fig. 2, as fractions of "left" responses as a function of gain. Two main features of the data can

[4]In our angle convention, 0° points to the subject's right, 180° to the left.

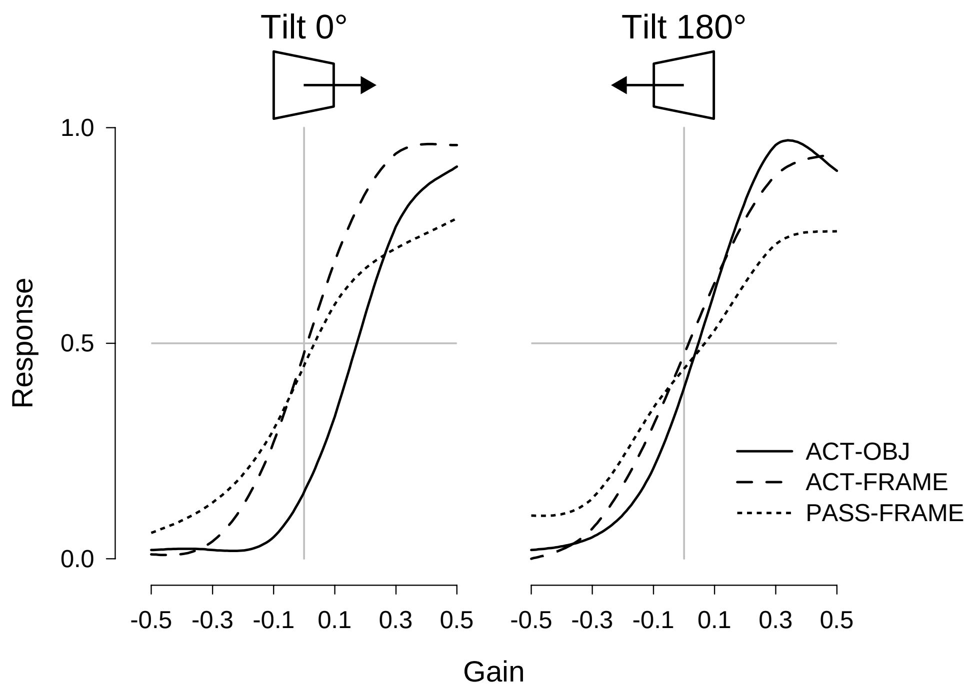

Figure 2: Raw data in Experiment 1, averaged for all subjects and smoothed. Fraction of "left" responses as a function of object rotation gain, for tilts 0° and 180°, in the ACT-OBJ, ACT-FRAME and PASS-FRAME conditions.

be noted by inspection of Fig. 2: the PASS-FRAME curves are wider than all others, implying that subjects were less precise in passive than active conditions; and the bias in the tilt 0° ACT-OBJ condition is farther from zero than in all other conditions.

In order to analyze these data quantitatively, we fitted them to logistic functions, as described in the General Methods section. Data from two conditions in each of two subjects were discarded, due to fitted bias falling outside the experimental range.[5] The average biases and widths, together with inter-subject standard errors, are shown in Fig. 3.

To test for possible effects of condition order, we performed a group (OBJ-FRAME, FRAME-OBJ) × condition (ACT-OBJ, ACT-FRAME, PASS-FRAME) × tilt ANOVA on both widths and biases, with group as a between-subject and condition and tilt as within-subject variables. Since neither the main effects of the group variable, nor any of its interactions proved significant, we discarded it from further analysis.

We begin with the widths—the relative precisions in the different conditions. Examining Fig. 3, we see no consistent effect of tilt, and not much difference between ACT-OBJ and ACT-FRAME conditions. On the other hand, the widths in the passive conditions are about twice as large as those in the active conditions. These effects were confirmed by a two-way ANOVA on condition (ACT-OBJ, ACT-FRAME, PASS-FRAME) × tilt variables, where the only significant effect was that of condition ($F_{2,14} = 11.2$). Evaluating specific contrasts, we find a significant difference between the ACT-FRAME and PASS-

[5]The omitted data was in the passive conditions, where two subjects had nearly flat response. Including this data only increased the numerical strength of the reported effects.

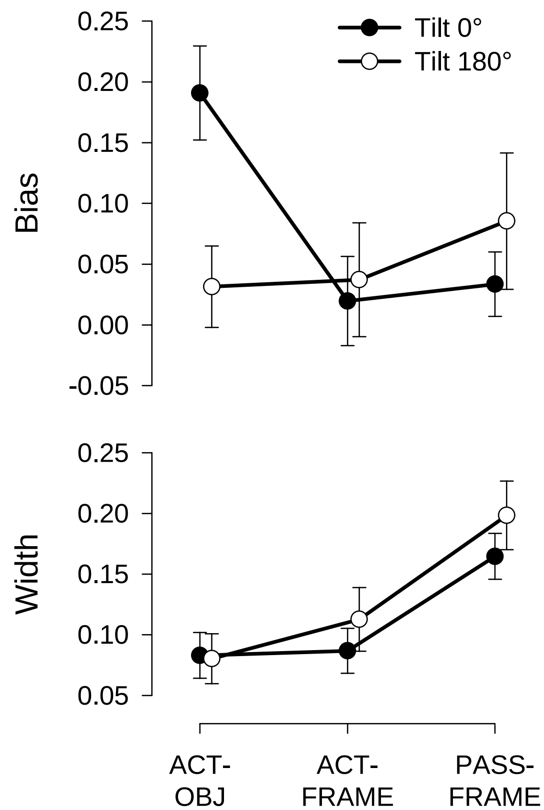

Figure 3: Fitted data from Experiment 1. Biases and widths (in units of gain) for tilts 0° and 180°, in the ACT-OBJ, ACT-FRAME and PASS-FRAME conditions. Fits were performed on individual subject data and resulting parameters averaged; error bars show between-subject standard errors.

FRAME conditions ($F_{1,7} = 13.7$), and no difference between the ACT-OBJ and ACT-FRAME conditions.

The story is rather different for biases. A condition in which the bias is zero corresponds to perfect accuracy. In active conditions, positive bias means either an underestimation of the subject's self-motion, or an underestimation of the visual effects due to self-motion (i.e., the rotation of a stationary surface relative to a moving observer). As can be seen from Fig. 3, the one condition where the bias differs from zero is ACT-OBJ, tilt 180°. A t test shows that bias in this condition is significantly positive ($t_9 = 4.94$, 2-tailed), while none of the other biases differs significantly from zero. The ACT-OBJ condition is also the only one where there is a significant effect of tilt ($t_9 = 3.83$, 2-tailed).

3.3 Discussion

These results show that both retinal and extraretinal criteria can be used to make stationarity judgments. On one hand, the ERC can be used without additional information: otherwise, the precision in ACT-OBJ would be worse than in ACT-FRAME, which it is not. On the other hand, the RC can also applied in isolation, since performance in PASS-FRAME is better than chance. The ERC (the only criterion available in the ACT-OBJ condition) yields stationarity judgements that are precise but, in the tilt 0° case, not very accurate (i.e., subject to systematic error). Retinal information (PASS-FRAME), on the other hand, yields judgements that are accurate but imprecise. Finally, when both sources of information are present (ACT-FRAME), performance reflects the best feature of each separate criterion: the high precision of ERC is maintained, as well as the high accuracy of the RC.

A puzzling feature of our data is the pattern of accuracy in the ACT-OBJ condition: at tilt 180°, the responses are accurate (i.e., mean bias does not differ from 0), but at tilt 0° there is a statistically significant positive bias. This difference is quite robust: in 9 out of 10 subjects the bias is greater (more positive) for tilt 0° than for tilt 180° in ACT-OBJ. One possible explanation could be that the estimate of self-motion coming from extraretinal signals underestimated true displacement. This explanation has been used to account for the Filehne illusion, for instance. An underestimate of self-motion would produce a positive bias, and could therefore explain the tilt 0° case in ACT-OBJ—but not the tilt 180° case, where the mean bias was no different from zero. Therefore, extraretinal underestimation of self-motion cannot account for our results. The phenomenon might have something to do with the way slant is perceived, which, at least in the passive observer, is subject to gross inaccuracy (Gibson, 1966). In the tilt 0° case slant relative to the observer increases during observer movement, while the opposite is true of the tilt 180° case. Little is known about how *changes* of slant are perceived, but it is conceivable that a net change in slant is overestimated during a slant increase, as compared to a slant decrease. In the ACT-FRAME and PASS-FRAME conditions, on the other hand, an error in slant estimation for the test surface could be canceled by a similar error for the background, eliminating any bias.

We may compare our results for biases and widths to those found by Wallach, Stanton and Becker, who also studied conditions similar to our ACT-OBJ and ACT-FRAME (Wallach, Stanton, and Becker, 1974). Their task was to indicate whether the object (which also rotated about its center as the observer moved, with an angular speed gain as in our experiment) appeared stationary or not, and they therefore measured gain thresholds at which subjects perceived motion, either in the same direction as their own rotation (corresponding to our positive gains), or against their rotation (as for our negative gains); these thresholds could be thought of as roughly analogous to our widths. The main result of Wallach, Stanton, and Becker (1974) is that the no-motion range did not depend on whether the test object was seen alone or against a stationary background. This is analogous to our finding that the widths in ACT-OBJ did not differ from those in ACT-FRAME. Furthermore, Wallach et al.'s data show that gain thresholds for with-subject rotations are quantitatively greater than against-subject thresholds (although this is not commented on, nor subjected to a statistical test). This result is in agreement with our finding that biases are either positive or zero, since in Wallach et al.'s experiments subjects viewed non-planar objects that had surfaces that corresponded both to tilts $0°$ and $180°$, on each trial. Finally, Wallach et al. found thresholds that are numerically much larger than our widths, by a factor of 3 or 4. This may be due to their very small stimuli ($2 - 3°$), as compared to ours (about $35°$). Another possible explanation is that for moderate gains, observers might not have a conscious perception of object motion (and therefore fall below threshold for Wallach et al.), but nevertheless discriminate the *direction* of motion much better than chance, and therefore be above our threshold.

In the first experiment, we have thus shown that stationarity judgments sometimes benefit in accuracy from retinal background information, but can be performed with higher precision based on extraretinal cues alone as compared to retinal cues alone. We thus have a hint that when both criteria are available (ACT-FRAME), both are used to some extent. However, we have no quantitative evidence about the relative weights of the two criteria. In Experiment 2 we address this question by using a cue-conflict paradigm.

4 Experiment 2

To determine the extent to which retinal and extraretinal criteria for stationarity are used when *both* are available, we performed a second experiment in which the two yielded conflicting results. Experiment 2 was similar to the ACT-FRAME condition of Experiment 1, namely the subject performed a movement around a scene where there was both a visual background and a foreground object. The foreground object rotated about the fixation point with a gain γ relative to the subject's rotation about the same point. In this experiment, however, the background was no longer stationary, but rotated with a gain γ_b with respect to the subject's movement about the same center as the foreground object, as shown in Figure 4. Subjects were told that the background would sometimes also rotate, and were instructed to ignore this rotation, basing their judgments of the foreground object's stationarity on earth-fixed criteria.

If subjects used purely retinal criteria—i.e., if they took the background as stationary, and judged stationarity of foreground objects relative to the background—all judgments would be based on the relative motion of the foreground to the background, namely $(\gamma - \gamma_b)\theta$. Thus the bias, γ_0, as

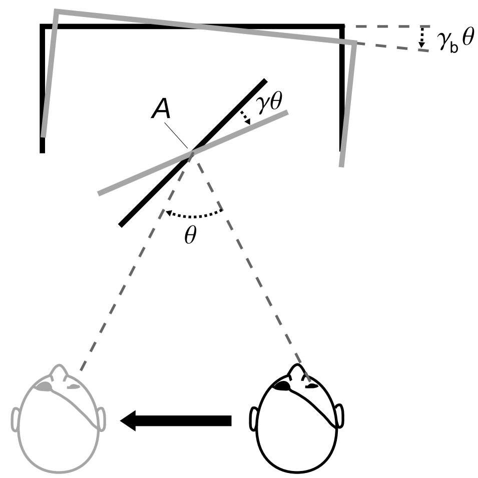

Figure 4: Simulated geometry and subject movement in Experiment 2. The subject's motion was translated into a rotation by angle θ about the stimulus center, marked A. The foreground object rotated by angle $\gamma\theta$ about the same axis (as in Experiment 1), while the background (not shown to scale) also rotated by angle $\gamma_b\theta$ (all rotations about point A).

a function of background motion, γ_b should increase linearly with a slope of 1. If, on the other hand, stationarity judgments were based purely on extraretinal criteria, subjects would be able to ignore background motion (as they were instructed to do), and bias would not depend at all on γ_b. We shall call these two extremes the retinal and extraretinal limits, respectively.

Experiment 2 bears comparison to a classic experiment that also pits retinal against extraretinal criteria: the rod-and-frame test (Witkin and Asch, 1948). In this experimental paradigm the subject adjusts a line so that it appears vertical, i.e., parallel to the direction of gravity. The line is surrounded by a square frame, which may itself be tilted. Other than the rod and frame, the experiment is conducted in darkness. To give the objective vertical, the subject must make use of extraretinal cues, namely vestibular information. On the other hand, the subject could also use retinal cues: in other words, he or she could assume that the vertical and horizontal directions are aligned with the frame. Most if not all subjects are at least partly influenced by the retinal cues—the subjective vertical is rotated from its true orientation in the direction of the frame. The amount of this retinal cue influence is called 'field dependence,' and is a repeatable measure for a given subject. The parallel between the rod-and-frame and our paradigm, though not exact (our experiment involves depth perception and motion, the rod-and-frame is static and 2D) was sufficiently tempting that we also administered the rod-frame-test to all subjects in Experiment 2, in order to see whether the use of retinal as opposed to extraretinal criteria was correlated to field dependence.[6]

4.1 Methods

The techniques used are described in the General Methods section. The slant and tilt of the surface were always 45° and 180°, respectively. The experimental design was factorial, with the foreground object gain $\gamma = -0.5$, -0.3, -0.1, 0.1, 0.3, or 0.5, and background gain $\gamma_b = -0.3$, -0.1, 0.0, 0.1, or 0.3. Each condition was repeated 10 times, for a total of 300 trials per subject, divided into two blocks of 150. The order of the trials was random.

Preceding the first block, the naive subjects were given two practice blocks of 20 trials each. In the first, both γ and γ_b were set to 0, and subjects

[6]The proximal cause for administering the rod-and-frame came from a pilot study for Experiment 2, which showed large inter-individual variations in criterion weights, reminiscent of variation in field dependence.

were told that both foreground and background would be stationary as they themselves moved, and that "the only reason why the image you see changes is because of your changing viewpoint." At the end of the first practice block, all subjects agreed that the scene had appeared stationary in an earth-fixed reference frame. In a second practice block, the foreground object gain γ was set to either 0.5 or -0.5, while the background remained stationary ($\gamma_b = 0$). The subjects performed the left/right discrimination task, and were given occasional feedback. After the first few trials, none of the subjects had any difficulty. Before the start of the main experimental blocks, subjects were told that the background would sometimes also rotate, and instructed to perform the left/right discrimination in an earth-fixed reference frame, i.e. that they should ignore any background rotation.

Nine subjects (6 men and 3 women, ages $25 - 33$ years, mean age 29), volunteered to participate in the experiment, including one of the authors and eight others who were naive to the experimental goals and hypotheses. Three of the subjects had also participated in Experiment 1.

Prior to the main experiment, a rod-and-frame test was administered to subjects in order to measure their degree of 'field-dependence' (Witkin and Asch, 1948). The stimulus was viewed binocularly at about 80 cm by the standing subject, with its center approximately at eye level. The stimulus consisted of a rod (length 35 cm, about 25°) whose center coincided with that of a square frame (length 70 cm, about 47°). All stimuli were drawn as thin, white and gray antialiased lines, with a black background. The frame was tilted by $\pm 10°$ or $\pm 20°$ from the upright, and the rod was presented with a random initial orientation between 10° and 30° or between $-10°$ and

$-30°$. The subject's task was to adjust the orientation of the rod so that it appeared vertical; subjects were explicitly told to ignore the orientation of the frame. The orientation of the rod was adjusted by pressing one of four buttons, which rotated the rod about its center by $\pm 5°$ or $\pm 0.5°$; a fifth button was used to validate the response. Each frame orientation was repeated 4 times for a total of 16 trials.

In the main experiment, biases and widths were calculated, as described in the General Methods section, for each subject for each value of background gain γ_b. The biases as a function of background gain, $\gamma_0(\gamma_b)$ were then linearly regressed against γ_b; as described above, the slope, which we shall call background dependence (BD), was expected to lie between 0 (the 'extraretinal limit') and 1 (the 'retinal limit'), and thus measured the relative contribution of retinal criteria to stationarity judgments. In the rod-and-frame, the mean rod inclination was linearly regressed against frame inclination in individual subject data; the resulting slope is a measure of visual field dependence (FD) for the subject.

4.2 Results

Fitted biases and widths are shown in Figure 5 as a function of background gain. The $\gamma_b = 0$ condition was the same as the ACT-FRAME, tilt 180° condition of Experiment 1; as can be seen by comparing Figures 3 and 5, the biases and widths did not differ significantly in the two experiments.

In order to calculate the background dependence (BD) for each subject, we performed a linear regression of biases against background gain γ_b. All correlations, which had r^2 between 0.85 and 0.98, were positive and statis-

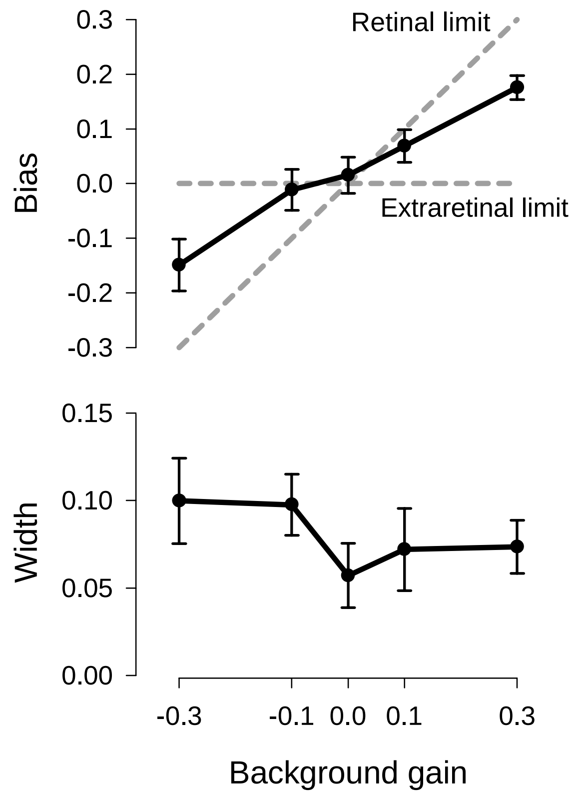

Figure 5: Biases and widths in Experiment 2, as a function of background motion. Error bars show between-subject standard errors, and the dashed lines in the bias plot show predictions based on pure retinal and extraretinal strategies.

tically significant. The BDs (the regression slopes) ranged between 0.24 and 0.78, with mean 0.49 and standard deviation 0.17.

To test the robustness the above procedure for calculating BD, we performed the calculation in an alternate way. We fitted all the data from each subject to a multivariate logistic function, $1/(1+e^{-z})$, with $z = a + b\gamma + c\gamma_b$. Calculated in this way, the BD is the relative weight of background to foreground motion, c/b. These logit BDs have a large and significant correlation with the BDs calculated using the other method ($r^2 = 0.94$, $t_9 = 10.7$); they range from 0.21 to 0.75, with a mean of 0.48 and a standard deviation of 0.17.

Having verified our procedure for calculating the BDs, we return to calculations using the original method. Does the mean BD agree with either the retinal (BD=1) or the extraretinal (BD=0) limit? The Gaussian 95% confidence interval for mean BD is 0.36 to 0.63. Thus, a two-tailed t test shows that the mean of the distribution is significantly greater than 0 ($t_8 = 8.62$) and less than 1 ($t_8 = 8.85$). It may be objected that a t test is inappropriate, given that we do not know if the underlying (population) distribution of BD is Gaussian. The weakest non-parametric test (i.e., the one that makes the fewest assumptions about the distribution), the sign test, also shows that the mean BD is significantly greater than 0 and less than 1 ($z = 2.67$ for both).[7]

Based on the rod-and-frame test, we also calculated a field dependence (FD) for each subject, as described in the Methods section. The FDs ranged

[7] In some ways, this 'test' is trivial, given that all of our samples lie between 0 and 1. Given n samples that are *all* greater than x, the sign test always yields $p < 0.05$ when $n \geq 6$.

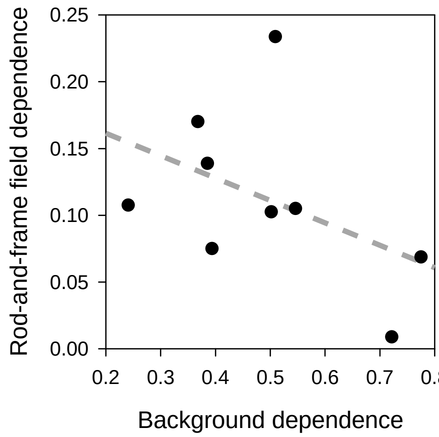

Figure 6: Two measures of visual field dependence, FD and BD, for each of the nine subjects in Experiment 2. The measures are anticorrelated, but not significantly.

between 0.009 and 0.234, with mean 0.112 and standard deviation 0.064. At the frame angles that we used (10° and 20°, these values are typical of past results on the rod-and-frame test (see, e.g., (Beh, Wenderoth, and Purcell, 1971)). Figure 6 shows the FD for each of the nine subjects, plotted against the BD calculated in the main experiment. The FDs and BDs were *anti*correlated, with $r = -0.449$ and a slope of -0.168. However, this anticorrelation was not significantly different from zero ($t_9 = 1.33$, $p = 0.22$). The 95% confidence limits on the coefficient of correlation are $-0.858 < r < 0.306$; thus at 95% confidence, we can exclude a positive correlation that has $r^2 > 0.1$.

4.3 Discussion

With this experiment we have shown that when both retinal and extraretinal criteria for stationarity are available, subjects in fact make use of both, with approximately equal weights. Given the results of Experiment 1—that stationarity judgments may be based on either criterion alone, but that the RC is accurate but imprecise, that the ERC alone is precise but sometimes inaccurate, and that the two in combination are both accurate and precise—using both criteria would therefore seem to be an optimal strategy.

We have also found large inter-individual variations in the weights assigned to the two stationarity criteria: in our nine subjects, the BDs ranged from 0.24 (strongly extraretinal) to 0.78 (strongly retinal). Given the analogy between our paradigm and the rod-and-frame test, one could expect that these variations would be correlated with the rod-and-frame field dependence measure; instead, we found an *anti*correlation ($r^2 = 0.202$)—but one that is not statistically significant.

5 General discussion

In the first experiment, we have shown that the extraretinal criterion is sufficient for making precise—though in some cases inaccurate—judgements of stationarity of 3D surfaces. This inaccuracy is reminiscent of the situation two-dimensional stationarity judgements in the presence of eye rotation, where, as in the Filehne illusion, self-motion is systematically underestimated. The *selective* character of the inaccuracy that we have found—it is only present for surfaces with increasing, not decreasing, slant—cannot be accounted for by an underestimation of slant, and remains to be explained.

By itself, the retinal criterion on the other hand does not seem to present any systematic bias, but instead suffers from another problem: low precision. When both the retinal and extraretinal criteria are available—as in the ACT-FRAME condition of Experiment 1, and as would be the case in everyday settings—the resulting stationarity judgements have the high precision of the ERC, and, as is the case for the RC, do not show any systematic bias.

Thus both the extraretinal and retinal criteria would seem to be necessary for stationarity judgements that are accurate and precise. We studied their relative weight in Experiment 2, by using stimuli where the two criteria yielded slightly different results. All subjects showed evidence for a mixture of the two criteria, but relative weights varied widely, ranging from subjects that rely mainly on the ERC (little visual capture) to those that rely mainly on the RC (high visual capture). It could have been supposed that the extent to which subjects rely on the RC reflects their 'field dependence'. However, this does not seem to be the case: when field dependence was measured by the standard rod-and-frame test, a negative (though not significant) correlation was found with reliance on the RC.

How are the two stationarity criteria combined? One model (resembling the 'weak fusion' of Landy et al. (1995)) would be for the extraretinal and retinal criteria to be applied independently, and their outputs to be combined linearly for a final stationarity estimate. Our data do not support this, since a linear combination of a precise and inaccurate estimator together with an imprecise and accurate estimator would not produce one that is both precise and accurate. The retinal and extraretinal criteria would thus seem to be combined in a more complex way.

In order to use the extraretinal criterion, an observer moving past a given ('nodal') point on a surface must determine the equivalent rotation angle about that point (e.g., angle θ in Figure 1a). This angle may be inferred from several different sources of information. If the observer fixes the nodal point and the head undergoes a pure translation, the equivalent angle is given by the rotation of the eyes with respect to the head. If the head turns as well as translates, the equivalent angle is given by the rotation of the eyes in *space*, that is the rotation of the eyes with respect to the head minus the rotation of the head in space. In the most complex—but also the most realistic—case, the observer may have fixated an arbitrary point on the surface, while his or head simultaneously rotated and translated in space. In that case, the equivalent rotation angle depends on the retinal motion of the nodal point, the rotation of the eyes with respect to the head, and on the rotation of the head in space. Alternatively, the equivalent rotation angle can be calculated from absolute distance to the nodal point (which would be known in ordinary, binocular settings) together with total head translation (Gogel and Tietz, 1973).

Since extraretinal information does play a role in the perception of stationarity, it is tempting to speculate on its origin. This information can be classified into one of two types, *primary* and *secondary*. Primary information is due to the active observer's self-generation of movement, hence premotor and motor brain activity with it efferent copy (von Holst and Mittelstaedt, 1950; Sperry, 1950). Secondary information is due to the feedback in the moving observer, such as vestibular and somatosensory afferents. It is interesting to consider, in this light, the absence of a strong positive correlation between field dependence in the rod-and-frame test and background dependence in Experiment 2 (in fact, they are probably anticorrelated): if the origin of the extraretinal information were vestibular, we might expect a positive correlation between the two measures. This is a hint that primary movement information plays a more important role than secondary in the perception of spatial stationarity. We must be careful, however, since the rod-and-frame differs in more ways from our Experiment 2 in other ways as well, for instance in the absence of movement or depth. Passive displacement would be a more direct way of deciding whether primary or secondary information contributes to stationarity judgments.

References

Beh, H.C., P.M. Wenderoth, and A.T. Purcell (1971). The angular function of a rod-and-frame illusion. *Perception and Psychophysics* 9(4): 353–355.

Brenner, E. and A.V. Berg (1996). The special rule of distant structures in perceived object velocity. *Vision Research* 36(23): 3805–3814.

Chatziastros, A., D.W. Cunningham, and H.H. Bülthoff (2000). Can we be forced off the road by the visual motion of snowflakes? Immediate and longer-term responses to visual perturbations. *Perception* 29 (Supplement): 118.

Crowell, J.A., M.S. Banks, K.V. Shenoy, and R.A. Andersen (1998). Visual self-motion perception during head turns. *Nature Neuroscience* 1: 732–737.

Descartes, René (1664). *Le traité de l'homme.* Clerselier, Paris. Translated as *Treatise of man* by T.S. Hall, Harvard University Press, 1972.

Dichgans, J. and T. Brandt (1978). Visual-vestibular interaction: Effects on self motion perception and postular control. In Held, R., W. Leibowitz, and Teuber H.L., editors, *Handbook of sensory physiology*, Vol. VII, pp. 755–804. Springer, New York.

Dijkstra, T.M., V. Cornilleau-Pérès, C.C. Gielen, and J. Droulez (1995). Perception of three-dimensional shape from ego- and object-motion: comparison between small- and large-field stimuli. *Vision Research* 35(4): 453–462.

Duhamel, J.R., C.L. Colby, and M.E. Goldberg (1992). The updating of the representation of visual space in parietal cortex by intended eye movements. *Science* 255(5040): 90–92.

Duncker, K. (1929). Über induzierte bewegung. *Psychologische Forschung* 12: 180–259.

Filehne, W. (1922). Über das optische wahrnehmen von bewegungen. *Zeitschrift für Sinnephysiologie* 53: 134–145.

Gibson, J.J. (1950). *The perception of the visual world.* Houghton-Mifflin, Boston.

Gibson, J.J. (1966). *The senses considered as perceptual systems.* George Allen and Unwin, London.

Gogel, W.C. and J.D. Tietz (1973). Absolute motion parallax and the specific distance tendency. *Perception and Psychophysics* 13: 284–292.

Haarmeier, T. and P. Thier (1996). Modification of the filehne illusion by conditioning visual stimuli. *Vision Research* 36(5): 741–750.

Haarmeier, T., P. Thier, M. Repnow, and D. Petersen (1997). False perception of motion in a patient who cannot compensate for eye movements. *Nature* 389(6653): 849–852.

Hildreth, Ellen C. and Constance Royden (1998). Computing observer motion from optic flow. In Watanabe, Takeo, editor, *High-level motion processing*, pp. 269–293. MIT Press.

Hogervorst, M.A., A.M. Kappers, and J.J. Koenderink (1997). Monocular discrimination of rigidly and nonrigidly moving objects. *Perception and Psychophysics* 59(8): 1266–1279.

Jenkin, M., L.R. Harris, U. Jasiobedzka, D.C. Zikovitz, H. Jenkin, and J. Zacher (2000). Perceptual stability during yaw head movement. *Perception* 29 (Suppl.): 20.

Koenderink, J.J. (1986). Optic flow. *Vision Research* 26(1): 161–179.

Koffka, K. (1935). *Principles of Gestalt psychology.* Harcourt, New York.

Landy, M.S., L.T. Maloney, E.B. Johnston, and M. Young (1995). Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Research* 35(3): 389–412.

Mack, A. and E. Herman (1973). Position constancy during pursuit eye movement: An investigation of the filehne illusion. *Quarterly Journal of Experimental Psychology* 25: 71–84.

Matin, L., E. Picoult, J. Stevens, M. Edwards, and R. MacArthur (1982). Oculoparalytic illusion: Visual-field dependent spatial mislocalizations by humans partially paralyzed with curare. *Science* 216: 198–201.

Mittelstaedt, H. (1990). Basic solutions to the problem of head-centric visual localization. In Warren, R. and A. H. Wertheim, editors, *Perception and control of self-motion*, pp. 267–288. Laurence Erlbaum.

Panerai, F., S. Hanneton, J. Droulez, and V. Cornilleau-Pérès (1999). A 6-dof device to measure head movements in active vision experiments: Geometric modeling and metric accuracy. *Journal of Neuroscience Methods* 90(2): 97–106.

Rogers, S. and B.J. Rogers (1992). Visual and nonvisual information disambiguate surfaces specified by motion parallax. *Perception and Psychophysics* 52(4): 446–452.

Royden, C.S., J.A. Crowell, and M.S. Banks (1994). Estimating heading during eye movements. *Vision Research* 34(23): 3197–3214.

Sperry, R.W. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *Journal of Comparative and Physiological Psychology* 43: 482–489.

Stark, Lawrence and Bruce Bridgeman (1983). Role of corollary discharge in space constancy. *Perception and Psychophysics* 34(4): 371–380.

Stevens, J.K., R.C. Emerson, G.L. Gerstein, T. Kallos, G.R. Neufeld, C.W. Nichols, and A.C. Rosenquist (1976). Paralysis of the awake human: Visual perceptions. *Vision Research* 16(1): 93–98.

Todd, J.T. (1982). Visual information about rigid and non-rigid motion: a geometric analysis. *Journal of Experimental Psychology: Human Perception and Performance* 8(2): 238–252.

Ullman, S. (1979). *The interpretation of visual motion*. MIT Press, Cambridge, Mass.

van Damme, W.J. and W.A. Grind (1996). Non-visual information in structure-from-motion. *Vision Research* 36(19): 3119–3127.

von Helmholtz, Hermann (1867). *Handbuch der Physiologischen Optik*. Voss, Hamburg.

von Holst, E. and H. Mittelstaedt (1950). Das reafferenzprinzip. *Naturwissenschaften* 37: 464–476.

Wallach, H., J. Bacon, and P. Schulman (1978). Adaptation in motion perception: Alteration of induced motion. *Perception and Psychophysics* 24: 509–514.

Wallach, H. and J.H. Kravitz (1965). The measurement of the constancy of visual direction and of its adaptation. *Psychonomic Science* 2: 217–218.

Wallach, H., L. Stanton, and D. Becker (1974). The compensation for movement-produced changes in object orientation. *Perception and Psychophysics* 15: 339–343.

Warren, Jr, William H. (1998). The state of flow. In Watanabe, Takeo, editor, *High-level motion processing*, pp. 315–358. MIT Press.

Wertheim, Alexander H. (1987). Retinal and extraretinal information in movement perception: How to invert the filehne illusion. *Perception* 16(3): 299–308.

Wertheim, Alexander H. (1994). Motion perception during self-motion: The direct versus inferential controversy revisited. *Behavioral and Brain Sciences* 17(2): 293–355.

Wexler, Mark, Ivan Lamouret, and Jacques Droulez (2000). Stationarity: An allocentric criterion in the perception of 3d structure from motion. In preparation.

Wexler, Mark, Francesco Panerai, Ivan Lamouret, and Jacques Droulez (2000). Self-motion and the perception of stationary objects. *Nature* . In press.

Witkin, H.A. and S.E. Asch (1948). Studies in spatial orientation iv. further experiments on perception of the upright with displaced visual fields. *Journal of Experimental Psychology* 38: 762–782.