



Perception of plane orientation from self-generated and passively observed optic flow

Jeroen Ja van Boxtel, Mark Wexler, Jacques Droulez

► To cite this version:

Jeroen Ja van Boxtel, Mark Wexler, Jacques Droulez. Perception of plane orientation from self-generated and passively observed optic flow. 2002. hal-00000070

HAL Id: hal-00000070

<https://hal.science/hal-00000070>

Preprint submitted on 22 Nov 2002

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Perception of plane orientation from self-generated and passively observed optic flow

Jeroen J.A. van Boxtel, Mark Wexler* and Jacques Droulez
Laboratoire de Physiologie de la Perception et de l’Action
CNRS, Collège de France
11, pl. Marcelin Berthelot
75005 Paris, France

November 19, 2002

Abstract

We investigate the perception of 3D plane orientation—focusing on the perception of tilt—from optic flow generated by the observer’s movement around a simulated stationary object, and compare the performance to that of an immobile observer receiving a replay of the same optic flow. We find that perception of plane orientation is more precise in the active than in the immobile case. In particular, in the case of the immobile observer the presence of shear in optic flow drastically diminishes the precision of tilt perception, whereas in the active observer this decrease in performance is greatly reduced. Furthermore, perceived slant is better correlated with simulated slant in the active observer. We conclude with a discussion of possible systematic biases in tilt perception from optic flow, as well as of various theoretical explanations for our results.

1 Introduction

While moving through a three-dimensional (3D) environment, observers extract important visual characteristics of that world from the two-dimensional (2D) image flow on their retina. From this optic flow, they can sometimes reconstruct the original 3D layout with a fairly high level of accuracy, given its severe underdefined nature. Under certain conditions, immobile observers can also extract 3D structure and motion of objects from optic flow alone. The reconstruction of structure from motion (SfM) has for long time attracted much attention [von Helmholtz, 1867, Wallach and O’Connell, 1953, Rogers and Graham, 1979].

*Corresponding author. E-mail: wexler@ccr.jussieu.fr

And although it is well established that object shape can be recovered in SfM tasks, under which conditions to what extent, is still not fully known.

It has generally been assumed that, at least as far as the perception of 3D shape is concerned, SfM depends only on retinal input. Therefore, according to this assumption, an active observer moving around a stationary 3D object should perceive the shape of that object in the same way as an immobile observer receiving the same optic flow—which would be generated by an equal-and-opposite rigid motion of the object (e.g. [Wallach et al., 1974]). In fact, the hypothesis that SfM depends solely on optic flow, coupled with the ecological prevalence of observer motion through a largely stationary environment, was used by Wallach et al. [1974] as a justification for the *rigidity assumption* for immobile observers.

However, a number of studies have compared the SfM performance of actively moving observers with that of immobile observers receiving more-or-less accurate replays of the same optic flow, and have found differences between the two conditions. From these results we can conclude that the purely retinal theory of SfM cannot be the whole story. These studies fall into two groups.

In the first group of studies [Rogers and Rogers, 1992, Dijkstra et al., 1995, Wexler et al., 2001a,b], subjects were presented with stimuli that admitted a small number of different solutions. The frequencies with which the solutions from this discrete set were perceived was different in active and immobile conditions. Subjects in the experiments of Wexler et al. [2001b], for example, could perceive 3D surfaces based on either perspective or motion cues, which were in conflict. It was found that in the active condition subjects made use of motion cues more often than they did while immobile, despite receiving the same retinal stimulation in the two conditions. The results from this group of studies can be summarized by the *stationarity assumption*, namely a bias towards perceiving objects whose 3D motion is minimal in an allocentric or earth-fixed reference frame. The stationarity assumption will be discussed further below.

In the second group of studies [Rogers and Graham, 1979, Ono and Steinbach, 1990, Peh et al., 2002, Panerai et al., 2002], differences in SfM performance were found in active and immobile observers, but in tasks that involved the perception of absolute length (either distance or depth), which immobile observers could not in principle perform with metric accuracy. The performance differences between active and immobile conditions that have been found in these studies thus also provide evidence for extra-retinal contributions to depth perception.

The goal of the present study is to compare the *precision* of active and immobile observers in an SfM task that (1) can be done in both self-motion conditions (and in this way differs from the second group of studies cited above) and (2) where any active-immobile differences would not be due to different choices from amongst a discrete set of solutions (as in the first group of studies cited). In order to satisfy these two goals, we use a task in which the subject has to indicate the 3D inclination of a planar surface. The subject perceives this 3D information from optic flow that is either generated by active head motion around a stationary, virtual object (the ACT condition), or while remaining immobile but experiencing a replay of the same optic flow (the IMMOB condition).

1.1 Self-motion and the stationarity assumption

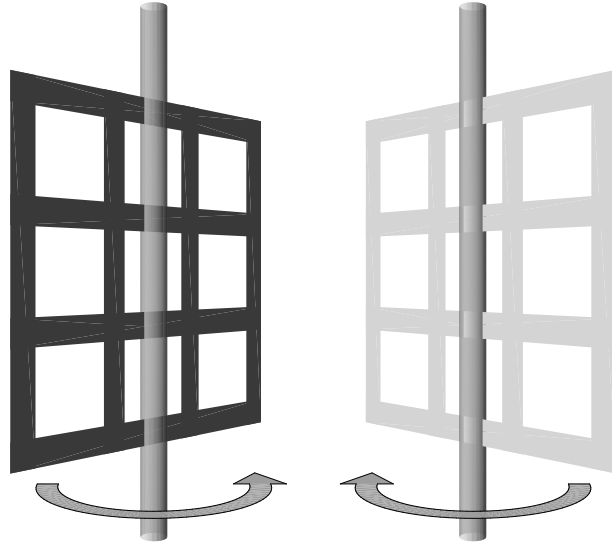
The advantage of incorporating information from extra-retinal sources is evident, as three-dimensional structure and motion are mixed in a non-linear way in optic flow. Confronted with an SfM task, an immobile observer has to simultaneously extract both structure and motion, and therefore must solve a complex, non-linear problem (see section 4.2.4). A moving observer, on the other hand, has additional extra-retinal information about motion, such as a copy of the motor command (in the case of voluntary motion) as well as proprioceptive information. This additional motion information can transform the SfM task into a linear problem, provided that relative motion between observer and object is due entirely to the observer’s self-motion—i.e., that the perceived object is stationary in an earth-fixed or allocentric reference frame. Since much of the optic flow that we experience is due to self-motion in a stationary environment, under many, or even most, circumstances this last provision is met.

Indeed, it has recently been shown that the visual system does make use of the heuristic assumption that objects are stationary in an allocentric frame—the stationarity assumption [Wexler et al., 2001a]—and that extra-retinal information about self-motion is used in this process [Wexler et al., 2001b]. One way in which we can see this is the reduction of *tilt reversals* in active observers as compared to immobile observers. A tilt reversal arises from the following symmetry: simultaneously adding 180° to the tilt of a rotating plane, and reversing its direction of rotation, results in approximately the same optic flow (see figure 1). Therefore, an immobile observer who has no *a priori* knowledge of the direction of rotation, is equally likely to perceive the simulated plane and its reversal (whose tilts differ by 180°). If the same optic flow is generated by an active observer moving about a stationary plane, however, the simulated and the reversed planes have very different motion in an allocentric frame: the simulated plane is stationary (by construction), while the reversed plane rotates with twice the speed of the observer. It is known that active observers perceive the reversed plane much less frequently than do immobile observers [Rogers and Rogers, 1992, Dijkstra et al., 1995], and it has recently been demonstrated that difference is due to the visual system’s making use of the stationarity assumption [Wexler et al., 2001a].

1.2 Surface perception from optic flow

Although many 3D objects have been used in experiments that probe SfM performance, one of the most simple ones is a plane. The orientation of a plane relative to the eye can be fully described by two variables, tilt and slant. Tilt is the orientation of the plane’s normal projected onto the fronto-parallel plane (in our convention, the direction of zero angle is to the subject’s right, with values increasing counter-clockwise, see figure 2). Slant is the angle, in 3D, between the normal and the direction perpendicular to the fronto-parallel.

For a small rotating planar object, slant and angular speed cannot be extracted independently from optic flow, which depends only on the product of its tangent and the



Insert animation here

Figure 1: The optic flow generated by the moving plane in our experiment (left) is approximately the same as the optic flow generated by a plane with its tilt rotated by 180° and reversed angular velocity (right). In the limit of small stimuli the difference between the optic flow generated by the two planes disappears, and they are equally likely to be seen by an immobile observer. The animation shows the motion, in an allocentric reference frame, of the simulated and reversed planes for active and immobile observers.

angular speed [Hoffman, 1982, Domini and Caudek, 1999].¹ Because slant is ambiguous in passive vision, observers’ judgment of this variable is not suitable as a measure to compare active and passive vision. Tilt, on the other hand, is theoretically well-defined [Ullman, 1979, Longuet-Higgins and Prazdny, 1980], even in immobile conditions, and serves this goal perfectly.

Whereas slant has received a great deal of attention in experimental studies (e.g., Braunstein [1968], Rogers and Graham [1979], Meese et al. [1995], Domini and Caudek [1999]), tilt has received relatively little. This is partly because tilt is theoretically well-defined and partly because, in most experimental studies, tilt is perceived without large errors, either random or systematic; this seems to be true in SfM tasks [Stevens, 1983, Domini and Caudek, 1999, Todd and Perotti, 1999, Norman et al., 1995], as well as in experiments where depth could be perceived from other depth cues like texture and shading [Norman et al., 1995]. The problem with most of these studies is that regression coefficients between perceived and simulated tilts are calculated over the entire range of tilt values, that is 360° . Consequently, it is hardly surprising that the slope of the regression line one finds is always near unity (remember that tilt is a circularly periodic variable, in contrast to the way slant is normally defined). Such an analysis passes over the possibility of systematic errors with a period of less than 360° , nor is it a very strong indicator of random errors (errors in precision), because these errors are not necessarily uniform over the entire range of tilt.

Cornilleau-Pérès et al. [2002] have found that the accuracy of tilt perception depended on the angle, in the image plane, between the normal and the direction of motion. Cornilleau-Pérès et al. [2002] called it the “winding angle”, which, for clarity, we shall call *shear angle* (see figure 2). Notice that the shearing motion—which is maximal when motion is orthogonal to the normal—diminishes as the angle between the axis of rotation and the tilt (i.e., projection of the normal) increases, and is absent when the two variables differ by 90° (which is the maximal difference between the two). Therefore we will define the shear angle to be 90° minus the smallest difference between the axis of rotation and the tilt. Using this definition the shear angle is zero when shear is zero, and the shear angle is 90° when shear is maximal (see figure 2). Cornilleau-Pérès et al. [2002] found that with an increasing shear angle the perception of tilt badly deteriorated, which is an interesting finding as tilt was believed to be easily and correctly found in SfM tasks.

No research has addressed the question of whether active vision increases the perceptual accuracy or precision, since the earlier studies concerned with tilt perception [Stevens, 1983, Domini and Caudek, 1999, Todd and Perotti, 1999, Norman et al., 1995, Cornilleau-Pérès et al., 2002] used objects presented only to immobile observers. Such results cannot automatically be extrapolated to active vision (see, for example, Wexler et al. [2001b], and the above-mentioned articles). Our study is the first to make a systematic comparison between the perceptual accuracy and precision during active and passive vision, while at

¹Strictly speaking the above ambiguities in tilt and slant hold only in first-order optic flow, i.e., in the limit of small objects or parallel projection, but they are approximately true for objects that subtend $5\text{-}10^\circ$, as in our experiments. In the case of perspective (polar) projection, parallel projections are approached in the limit of small objects.

the same time using a variable—tilt—which is suited for such a task.

2 Methods

2.1 Stimulus

In the reference frame used to describe the experiment, the xy -plane is co-planar with the monitor screen, with the x -axis pointing to the subject’s right, the y -axis upwards, the z -axis towards the subject, and the origin at the center of the monitor. Lengths will be expressed in centimeters.

The stimulus was a planar patch, inclined in depth. The patch was filled with a random-dot texture, with the dot distribution chosen so as to be uniform (on the average) in the projected image. This was done so as to remove texture cues to depth as much as possible: see figure 3. The 200 dots were chosen so that their projections fell inside a circle of radius 5.25 cm in the image; therefore, the texture on the stimulus plane was an ellipse with a non-uniform distribution of dots. (At the approximate mean observation distance of 60 cm (see below), this resulted in a radial angular stimulus size of 5° .) During each moment (more precisely: display monitor frame) that the stimulus was visible, the texture elements were projected onto the screen using a perspective projection from the subject’s current eye position (measured by a head tracker; see below), and drawn as white pixels. The center of the stimulus lay in the xy -plane. The stimulus was centered at the point directly opposite the subject’s eye at the beginning of each trial (i.e., if the subject’s eye was at point (x, y, z) , the center of the stimulus was at $(x, y, 0)$).

The plane’s normal was $(\sin \sigma \cos \tau, \sin \sigma \sin \tau, \cos \sigma)$, where σ is the slant and τ is the tilt. Tilt varied from 0° to 345° in increments of 15° . Slant was 30° , 45° or 60° . A red fixation mark (a circle of size: 0.05 cm) was visible in the center of the stimulus during the entire duration of the trial. Other than the stimulus, nothing was visible, including the borders of the display monitor.

2.2 Motion

There was always relative motion between the subject and the stimulus. The SELF-MOTION parameter specified who, or what, was doing the moving. In the active condition (ACT), observers performed head movements around the simulated planar object, which remained stationary in an earth-fixed reference frame. In the immobile condition (IMMOB), observers did not move, while the virtual object whose projection generated the stimulus underwent rigid motion. The latter motion (composed of rotations about the fixation point, and scale changes) was such that optic flow in every IMMOB trial exactly replicated the observer-generated optic flow in a previous ACT trial. (Details are given in Appendix A.) In ACT, subjects moved either horizontally or vertically, which yielded two different axes of rotation.

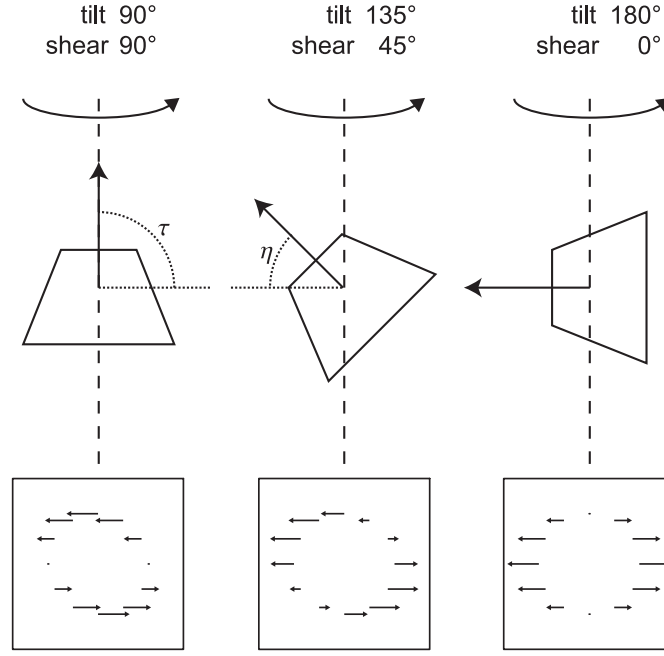


Figure 2: An illustration of the shear variable. *Upper Panels.* Three different orientations of a plane are depicted, with different values of tilt. The axis of rotation is considered to be vertical (which, in our experiment would be the case in the HORIZ motion condition). The left panel depicts how the tilt (τ) is defined, namely the orientation of the plane's normal (indicated by the arrow attached to the plane) projected onto the fronto-parallel plane. The middle panel depicts the definition of the shear angle. The shear angle (η) is 90° minus the smallest difference between the axis of rotation and the tilt. The actual values of the tilt, slant and shear (or shear angle) are presented above the panels. With this axis of rotation the shear and tilt have the same value in the tilt range 0° to 90° . *Lower Panels.* The approximate optic flow associated with the conditions drawn in the upper panels are shown. The left panel shows the optic flow with 90° shear angle, the maximal value (pure shearing motion). The right panel shows the optic flow when the shear angle is equal to zero, its minimal value. The middle panel shows the flow for an intermediate shear angle.

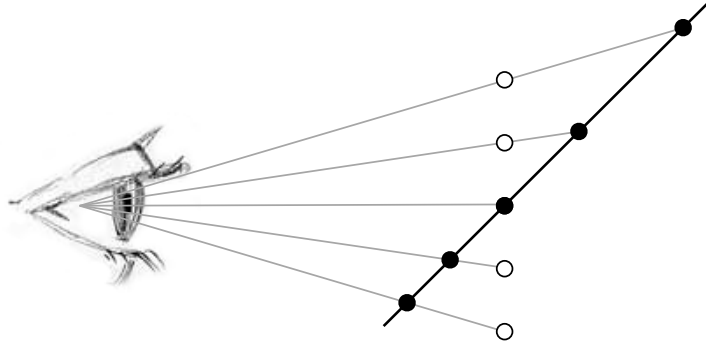


Figure 3: A schematic diagram of texture in our stimuli. The goal is to remove texture (i.e., non-motion) cues to depth as much as possible. We start with a uniform distribution of dots in the image plane (the white circles). These are then projected onto the inclined stimulus plane (the black circles). The distribution of the black circles is therefore non-uniform in the stimulus plane. With only small movements of the object or observer, the distribution of texture elements in the image remains nearly uniform.

2.3 Procedure

The experiment was performed in monocular viewing conditions using the subject’s dominant eye, with the other eye covered. We will refer to the position of the center of the dominant eye (as measured by the head tracker—see below) simply as the “subject’s position”. For a trial to begin, the subject’s position had to be inside a cube with sides of length 10 cm, centered on the point $(0, 0, 60)$ (so that the stimulus did not surpass the edges of the monitor screen).

In the beginning of every ACT trial, the subject was verbally cued to begin moving his head in one of four directions: right, up, left, or down (from the subject’s point of view). Initial motion cycled on every subsequent trial through these four directions. The DIRECTION variable grouped trials by direction of motion: left and right trials will be called HORIZ, and up and down trials VERT. Note that, in terms of relative rotation between the subject and the plane, HORIZ trials resulted in a vertical axis of rotation, while VERT trials resulted in a vertical axis. We used both horizontal and vertical motion to decorrelate shear and tilt.

Motion continued until displacement along the required direction reached 3 cm, at which point a beep was heard. This was the signal to reverse the direction of motion. When, after reversing direction, the subject’s position reached -3 cm along the direction of motion, another beep was sounded, etc. In this way, we produced oscillatory motion in a given direction. The subject performed this oscillatory motion for 2.5 cycles in each trial. During the first half-cycle, only the fixation mark was visible, while the stimulus appeared during the next 2 cycles. Following the 2.5 cycles, the stimulus disappeared, and was replaced by a response probe. This was the subject’s signal to stop moving the head.

To control variability in motion trajectories, we implemented some additional restric-

tions on the subject’s motion. The amplitude was controlled by aborting the trial if displacement along the required motion direction (x - or y -axis in HORIZ and VERT trials, respectively) exceeded 6 cm from the central point. To make sure that motion was primarily in the required direction, at the end of each trial we calculated the RMS of the subject’s motion in that direction, normalized by the RMS of the motion in the two perpendicular directions; if this ratio exceeded 0.5, the trial was restarted. Furthermore, a trial was restarted when the duration of the visible stimulus was less than 2 s or greater than 5 s.

In the IMMOB condition, the subject’s motion was also measured using the head tracker. The same condition was applied as in ACT for a trial to begin. The subject was instructed to remain still for the remainder of the trial. Nevertheless, any motion that the subject produced was taken into account in reproducing the optic flow from the corresponding ACT trial, as described in the Appendix. Furthermore, all other cues from the ACT trial (the verbal cue indicating the previous initial direction of motion, the beeps to control the subject’s movement) were replayed during the IMMOB trial.

Following the presentation of the stimulus in each trial, subjects had to indicate the perceived inclination of the plane by adjusting a visual probe with a joystick. The probe was a square grid, subdivided into 6×6 squares (each 1.75 cm wide, total size 10° if the probe had zero slant and the subject was 60 cm from the screen). The orientation of the grid texture was random on each trial, in order to prevent any reliable 2D cues. (Probe texture orientation in IMMOB trials was the same as that in corresponding ACT trials.) The slant of the probe was proportional to the polar angle of the joystick shaft, while the tilt was equal to the joystick azimuth. The probe had a maximum slant of 80° .

We used a factorial design. Each subject performed 576 trials: 2 SELF-MOTION conditions, 3 slant values, 24 tilt values, and 4 directions of initial motion. The experiment was performed in 3 ACT and 3 IMMOB blocks, with each ACT block followed by the corresponding IMMOB block. The order of ACT trials within each block was random. The IMMOB blocks reproduced each trial, in the same order, from the previous ACT block. Before the experiment started, subjects were given 2 practice blocks, one active and one immobile.

2.4 Apparatus

The translational eye displacements of the subject were measured by a mechanical head tracker [Panerai et al., 1999], which has submillimeter within-trial precision. Sampling of the head tracker was at the same frequency as the display monitor, 96 Hz. The latency exhibited by the tracker was lower than the sample interval. A Pentium II 400 MHz computer both sampled the tracker (using a National Instruments PCI-6602 card) and controlled the stimulus display (Sony GDM-F500 CRT monitor with 1600×1200 pixels on a 40.2×29.6 cm screen, driven by a Matrox G400 video card). The resolution was about 1.4 arcmin/pixel at a distance of 60 cm. A Microsoft Sidewinder Precision Pro digital joystick was used to direct the probe. Subjects viewed the stimulus monocularly with their dominant eye, the other being covered by an opaque patch. The experiment was performed in a dark room. To prevent anything other than the stimulus from being seen, the observers wore a pair of ski goggles.

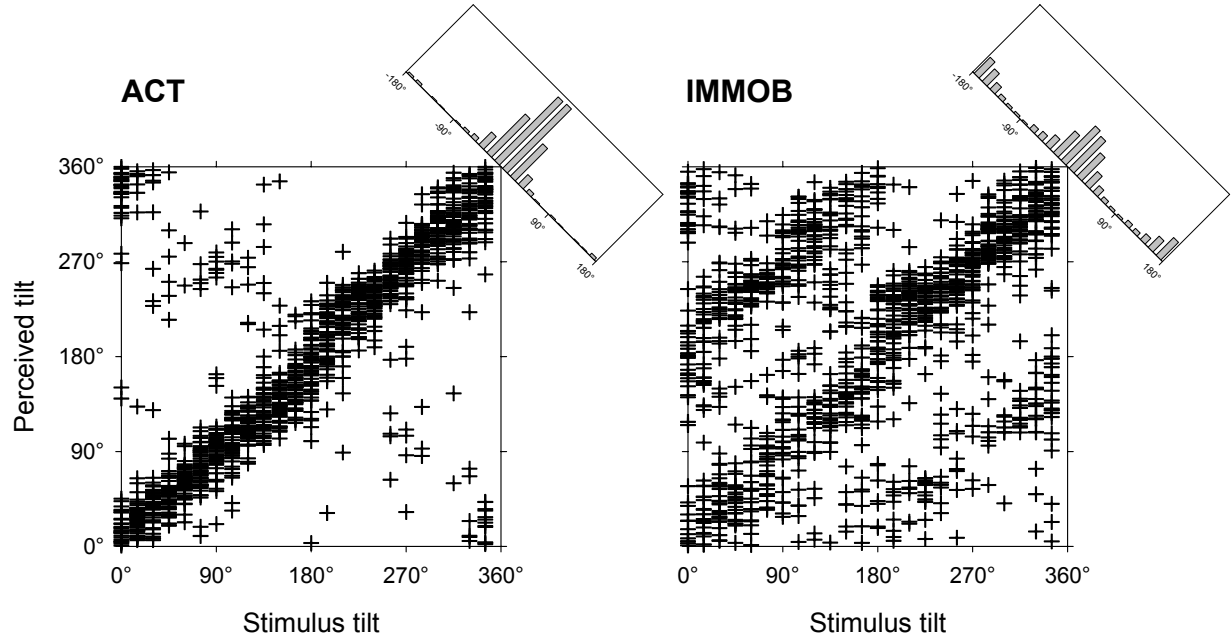


Figure 4: Response vs. simulated tilt for all subjects in the ACT and IMMOB conditions. The histograms show the distributions of the differences Δ_τ between response and simulated tilt.

2.5 Subjects

Five subjects naive as to the purpose of the experiment, aged between 20 and 25 years (2 men, 3 women), participated. All had normal or corrected-to-normal vision. They were paid on an hourly basis.

3 Results

It was found that trials where the initial motion was to the left gave the same results as trials where the initial direction was to the right, and likewise for up and down initial motion. This variable is therefore not discussed in the rest of this paper. Neither any differences were found between the first, second and third sessions. We therefore collapsed all data across these variables.

3.1 Tilt errors

In figure 4 we plot the response versus the simulated tilt. We define the quantity Δ_τ as the smallest angular difference between the response and simulated tilt, ranging from -180° to 180° . The histograms in figure 4 show the distributions of Δ_τ . The figure shows that, especially in the immobile (IMMOB) condition, in many trials Δ_τ was close to $\pm 180^\circ$,

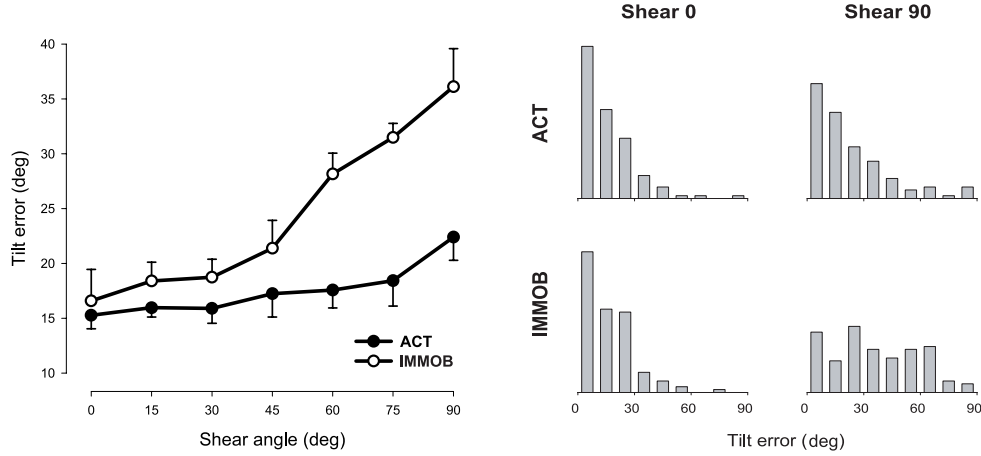


Figure 5: The influence of the shear angle on tilt error E_τ . In IMMOB, a clear increase of tilt error is observed with increasing values of the shear angle. The increase is present but significantly smaller in ACT. In the right panel four histograms show the tilt error distribution in both ACT as IMMOB, for 0° and 90° of shear angle. For 0° of shear, tilt errors in both ACT and IMMOB are small. For 90° of shear angle, errors have increased, especially for IMMOB, where performance is near chance level. Error bars represent between-subject standard errors.

which corresponds to the perception of a reversal (see Introduction). Since the optic flow is ambiguous—we really have two simulated tilts, which differ by 180° —we introduce an absolute-value tilt error measure, E_τ , that corrects for possible reversals:

$$E_\tau = \begin{cases} |\Delta_\tau| & \text{if } -90^\circ < \Delta_\tau < 90^\circ \\ 180^\circ - |\Delta_\tau| & \text{otherwise} \end{cases} \quad (1)$$

As defined, E_τ ranges from 0° to 90° . In the rest of this article we will refer to E_τ as the ‘tilt error’.

Figure 5 shows the dependence of the tilt error on the shear angle. There is a clear difference between the ACT and IMMOB conditions. In ACT the mean tilt error is 17.3° , in the immobile condition it is 24.1° , and this difference is significant ($p < 0.01$). A higher tilt error in IMMOB than in ACT was observed in all subjects. However, the average error is not fully informative, as there is a large effect of shear angle.

A SELF-MOTION (ACT, IMMOB) \times DIRECTION (VERT, HORIZ) \times slant \times shear angle ANOVA with tilt error as a dependent variable, showed that shear had a significant effect on the precision of tilt perception ($p < 10^{-4}$). Further analysis showed that in both IMMOB and ACT, tilt error increased significantly with increasing shear (both $p < 10^{-3}$, 1-tailed t test). However, the tilt errors increased differently in the two conditions. The ANOVA showed a significant SELF-MOTION-shear angle interaction ($p < 10^{-4}$): the tilt error rose faster in the IMMOB than in the ACT condition. The magnitude of this effect can be demonstrated by a linear regression: the mean slope of the tilt error versus shear is 0.224

in IMMOB but only 0.067 in ACT; moreover, this slope is lower in ACT than in IMMOB in all subjects. The ANOVA further revealed that tilt errors decreased with increasing slant ($p < 10^{-3}$). Finally, the ANOVA revealed that there was a significant influence of AXIS, which we will return to in section 3.4.

3.2 Slant errors

Errors in slant perception were analyzed using a SELF-MOTION \times DIRECTION \times slant \times shear angle ANOVA with slant response as a dependent variable. This revealed a dependence on the simulated slant ($p < 10^{-4}$), although this dependence is rather weak (the slope of the linear regression is 0.21); however, all subjects showed a significant positive correlation between simulated and response slant in both SELF-MOTION conditions (all $p < 10^{-2}$). However, slant response was better correlated with simulated slant in ACT than IMMOB (mean slope 0.25 in ACT, 0.16 in IMMOB). Indeed, there was a significant SELF-MOTION-slant interaction ($p < 0.05$).

The shear angle, apart from its effect on tilt perception, also influences the ability of subjects to estimate slant: absolute slant error (i.e. the absolute difference between the response slant and the simulated slant) increases with increasing shear ($p < 0.05$).

3.3 Reversals

In the preceding analyses we neglected the occurrence of reversals. In this section we specifically look at the rate of reversals and at the effect reversals have on tilt and slant errors.

As stated above, we define tilt reversals as those trials in which the response tilt differed from simulated tilt by more than 90° . In IMMOB, reversals occurred on 35.3% of all trials, while they did so on only 4.4% in ACT (significant difference: $p < 10^{-4}$, z test for independent proportions). The rate of reversals in IMMOB is significantly different from 50% ($p < 10^{-4}$, z test for independent proportions).

As shown in figure 6, when reversals occurred, both tilt and slant errors increased significantly in ACT ($p < 0.05$), whereas they changed little in IMMOB. Nevertheless, even when errors were greater in reversal trials, the responses were not random: tilt responses in reversal trials were centered around reversed tilt, and almost absent in the region which could be interpreted as large deviations in the percept of the simulated tilt—see, for example, the histograms in figure 4. Absolute slant errors in IMMOB did not increase when reversals occurred, but a significant increase was seen in ACT ($p < 0.05$).

3.4 Anisotropies in plane perception

Although the distribution of simulated tilt was isotropic, the distribution of tilt responses was not. Here we present three separate analyses of this anisotropy. In addition, we found an anisotropy with respect to movement: slight but significant differences in responses were found between the HORIZ and VERT axis conditions.

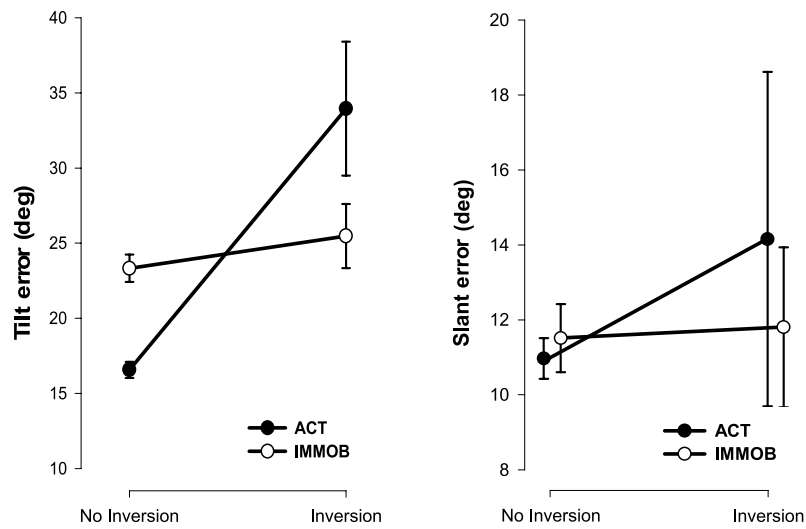


Figure 6: Reversals were found to have a significant effect on both tilt and slant perception. *Left panel.* In trials without reversals the tilt error was significantly smaller in ACT than in IMMOB. In reversal trials the errors in the IMMOB condition changed little, but the errors in ACT increased significantly and even became significantly greater than in IMMOB. *Right panel.* Slant errors were low in non-reversal trials for both ACT and IMMOB. For reversal trials the slant errors in IMMOB stayed at this level, but in the ACT condition the errors increased significantly, although there was a wide scatter between subjects. Error bars represent between-subject standard errors.

3.4.1 Directional biases in tilt responses

The first analysis of anisotropy that we carried out was a Rayleigh test on the tilt responses [Batschelet, 1981]. This test of directional anisotropy of a set of angles θ_i is based on averaging the two-dimensional unit vectors for each angle, $(\cos \theta_i, \sin \theta_i)$, and examining the length of this average vector. We carried out this test on the tilt responses, separately for each subject in the ACT and IMMOB conditions. Four out of five subjects showed a significant ($p < 0.05$) directional bias in IMMOB; in three of these subjects, the bias was between the angles 240° and 300° , i.e., ceiling-like surfaces whose normals are downwards. Only one subject had a significant bias in ACT, and it was directed upwards. Of the six remaining non-significant data points, four were directed downwards. If we group ACT and IMMOB trials, four out of five subjects showed a bias towards downward-orientated normals.

However, the above results confound two possible sources of anisotropy, namely tilt errors (corrected, up until now, for tilt reversals), and tilt reversals themselves. In the next two sections, we analyze these sources separately.

3.4.2 Systematic errors in tilt responses

So far, we have used the absolute-value tilt error, E_τ , which has confounded random with systematic errors in tilt. Here, we wish to examine systematic errors in tilt, which we correct for reversals, and therefore we define a new error measure,

$$S_\tau = \begin{cases} \Delta_\tau + 180^\circ & \text{if } \Delta_\tau < -90^\circ \\ \Delta_\tau & \text{if } -90^\circ < \Delta_\tau < 90^\circ \\ \Delta_\tau - 180^\circ & \text{if } \Delta_\tau > 90^\circ \end{cases} \quad (2)$$

where Δ_τ is the smallest angular difference between the response and simulated tilts, and the new, signed error measure S_τ ranges from -90° (clockwise errors) to $+90^\circ$ (anti-clockwise errors). Averaging S_τ for a given value of simulated tilt permits us to study any systematic bias that is present at that point, independently of reversals.

Indeed, systematic errors were present in our data—see figure 7. The effects are qualitatively similar in ACT and in IMMOB. The average of the systematic errors per tilt value over all subjects ranged from -12° to 18° in IMMOB and from -11° to 10° in ACT. On average the systematic error is 1.4 times greater in IMMOB than in ACT.

We carried out Rayleigh tests for bimodal distributions on the overall data, and for subjects individually. Of the two preferred directions, the one closest to 0° of tilt is given, the other angle differs by 180° .

The overall data showed a significant bias whose mean tilt is 84.9° in ACT ($p < 10^{-4}$, Bonferroni corrected) and 88.3° in IMMOB ($p < 10^{-4}$, Bonferroni corrected). The preferred directions of individual subjects were, 90.9° , 83.5° , 84.7° , 81.9° , 176.2° , in ACT, and 98.5° , 84.7° , 79.9° , 99.9° , 177.9° in IMMOB (keeping the order of the subjects the same). All 10 tests were significant at $p < 10^{-4}$ and remained so when Bonferroni corrected. These results show that four out of five subjects had a preference for approximately horizontally orientated planes, while one subject preferred vertically orientated ones.

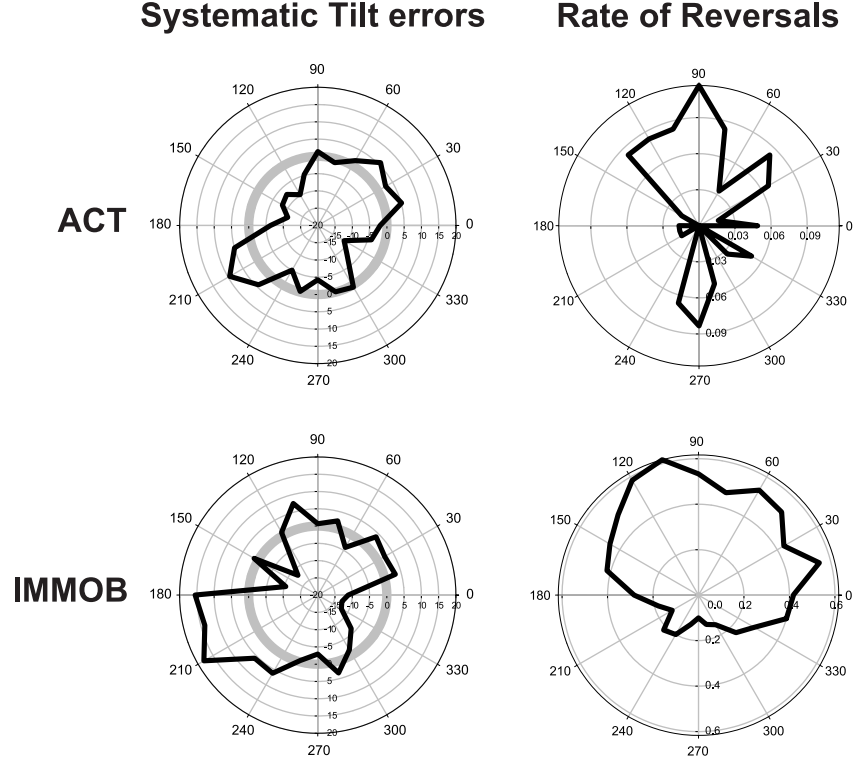


Figure 7: *Left panel.* The dependence of systematic deviations in tilt response S_τ (signed tilt errors, corrected for reversals) on simulated tilt. The thick gray line represents zero systematic bias. Anti-clockwise biases are given as positive values. The mean bias is towards 70.7° (and 250.7°) in ACT and 78.7° (and 258.7°) in IMMOB, i.e. roughly horizontal surfaces. *Right panel.* Rate of reversals versus the simulated tilt. In both ACT and IMMOB, reversals preferentially occur when the normal points upwards, generating a bias towards downward perceived normals (i.e. a ‘ceiling’-like plane). Note that the axis in IMMOB runs from 0 to about 0.6 and in ACT from 0 to about 0.12.

3.4.3 Reversal occurrence

The previous section showed that the outcome of the initial Rayleigh test was partly a result of significant systematic errors in the tilt perception. In this section we will look at the influence of reversals.

To investigate this we performed a Rayleigh test for unimodal data on the collection of trials which showed reversals, and averaged the unit vectors of the simulated tilt of these trials. When Bonferroni corrected, the data showed a significant effect in ACT (mean tilt 70.7° , $p < 10^{-4}$) and in IMMOB (78.7° , $p < 10^{-4}$). Both averaged vectors pointed upwards. A more conservative z -test for independent proportions, comparing the proportion of reversals in the upper hemi-field versus the lower hemi-field, showed a significantly higher amount of reversals in the upper hemi-field. This was true in ACT, as well as in IMMOB ($p < 0.01$ and $p < 10^{-4}$, respectively). There was no such difference between the left and right hemi-fields.

Both tests show that reversals occurred more often when the normal of the plane was directed upwards compared to downward, resulting in a bias towards seeing ‘ceiling’-like planes.

3.4.4 Anisotropy with respect to movement

The tilt error was significantly different in the two DIRECTION conditions, being greater in the VERT (26.8°) than in the HORIZ condition (22.5° , $p < 0.05$). Further analysis showed that the two curves in IMMOB were not significantly different, but the ones in ACT were. The actual difference between the two curves of the IMMOB conditions and the difference between the two ACT curves was, however, the same.

3.5 Analysis of movement trajectories

We analyzed the movement trajectories of the ACT condition, and studied kinematic quantities such as the maximum amplitude, the velocity and acceleration along all three axes. We divided the range of values each quantity subtended into several equally sized bins and looked if the data (tilt error, signed tilt error, and reversals) showed a dependence on the quantity in question. No such dependence was found.

Next, we compared the VERT versus the HORIZ condition to investigate the origin of the anisotropy with respect to movement. VERT trials showed larger amplitudes in the movement along the z -axis than HORIZ trials. The displacement along the x -axis during up/down movement was also greater than the displacement along the y -axis during left/right movement. We homogenized the trajectories *post hoc* by only considering trials whose movement amplitudes fell within a certain range. After this homogenization the anisotropy with respect to movement was, however, still present.

4 Discussion

4.1 Active vision is more precise than passive vision

We examined the perceptual precision in an SfM task using two dependent variables: tilt and slant. To compare active and immobile conditions, it is necessary that the dependent variable be well-defined and recoverable in both conditions. As other variables used in earlier research (e.g., depth and absolute distance) the precision of slant perception is for this reason not useful, as it is fully ambiguous in the IMMOB condition. The precision of tilt perception, however, does satisfy this criterion.

We find that tilt perception in ACT is more precise than in IMMOB, and thus demonstrate, for the first time, that active vision increases the precision of surface perception compared to passive vision. For optic flow with minimal shear, tilt precision is about equal in active and immobile conditions; however, as shear increases, precision falls off rapidly in IMMOB, while remaining almost constant in ACT (see figure 5).

4.2 In search for an explanation for the shear effects

Given its clear importance for surface perception, the shear variable has been very little studied, and its effect is not understood theoretically. In this sub-section, we explore several different ways to account for the effect of shear. However, we warn the reader from the outset that none of our models gives, at present, a satisfactory account. For readers wishing to skip the details, a summary of our arguments is given in section 4.2.5.

4.2.1 Change of surface normal during a trial

While the motion of the simulated surfaces in our experiment relative to the subject was the same in the ACT and IMMOB conditions, their motion in an earth-fixed or allocentric reference frame was different. Namely, the object rotated in IMMOB, while remaining motionless in ACT. During the object’s rotation in IMMOB, its normal—and therefore its slant and tilt, which we define here in an allocentric frame—changed. Could this allocentrically defined change account for the different effects of shear in the two conditions?

For small surface rotations in the IMMOB condition (first-order in rotation angle, α , a reasonable assumption for experiment where the average maximum angle was about 4°), the changes in slant and tilt are

$$\Delta\sigma \approx (\cos\eta)\alpha \tag{3}$$

$$\Delta\tau \approx \frac{\sin\eta}{\tan\sigma}\alpha, \tag{4}$$

where η is the shear angle. These equations contradict our results in the IMMOB condition in several different ways. First, the errors predicted in equations (3) and (4) are much too small: on the order of 4° for slant and 7° for tilt, as compared to our experimental results in the IMMOB condition of 12° and 40° , respectively. Second, equation (4) predicts

a milder dependence of tilt error on shear for higher slant in IMMOB; instead, averaging tilt error for $\eta = 0^\circ$ and 15° and subtracting from the average for $\eta = 75^\circ$ and 90° , we found tilt error differences of 14.6° , 18.7° , and 15.6° for slant 30° , 45° and 60° , respectively. Third, the $\sin \eta$ term in equation (4) would predict that the slope of the tilt error vs. shear curve approach zero as shear approaches 90° , which is not observed in our data. Fourth, equation (3) would predict that slant errors decrease with shear, whereas they significantly increase with shear.

Thus, the change in the allocentric normal in IMMOB cannot account for our results.

4.2.2 The stationarity assumption

It has recently been shown in our laboratory [Wexler et al., 2001a,b] that a new hypothesis is needed to account for structure-from-motion performance in moving observers: the visual systems makes the *stationarity assumption*, that is, prefers SfM solutions that minimize motion in an allocentric or earth-fixed reference frame. This bias has obvious computational and ecological advantages. If the SfM process is seen as iterative (e.g., as a search for a combination of 3D structure and motion that maximizes rigidity), one way that the stationarity assumption could be implemented could be as an initial condition: the visual system could use self-motion information to begin its search by considering motion that is null in an allocentric frame (which, if the observer is moving, would translate to non-zero motion relative to the observer). This type of temporal dependence is at least partly supported by the spatio-temporal coherence results of van Damme and van de Grind [1996].

Due to the stationarity assumption, the rate of reversals should be much smaller in ACT than in IMMOB: the simulated plane in ACT is stationary and the reversed plane is not, whereas the simulated and reversed planes are equally non-stationary in IMMOB [Wexler et al., 2001a]. We have indeed found this to be the case (see figure 4). (The fact that reversals occur in less than 50% of IMMOB trials indicates that the visual system takes into account second-order optic flow.)

The stationarity assumption would predict that solutions that really are stationary (as always, we mean stationary in an allocentric frame) will be perceived more precisely, since the initial 3D motion estimate will need much less refinement. This prediction is, in fact, borne out by some of our results. When there is no tilt reversal, the solution in ACT is stationary, while the solution in IMMOB rotates at the same speed, ω , as the subject did in the corresponding ACT trial (see animation).² Accordingly, we found that, in trials without reversals, tilt errors are smaller in ACT (16.6°) than in IMMOB (23.3°): see figure 6. When tilt reverses, in IMMOB the reversed solution rotates in the opposite direction but with the same speed, $-\omega$. In ACT, on the other hand, the reversed solution rotates at 2ω (see animation). Accordingly, we find IMMOB tilt errors about the same in reversed trials (25.5°), while in ACT they are about twice as high (34.0°) as in unreversed trials! This could mean that observers do not only prefer to see an allocentrically stationary object,

²This is only true, of course, if the subject estimates relative motion to the object, and self-motion in an allocentric frame, accurately. There is converging evidence that self-motion is underestimated, but only by about 30 – 40% in actively moving subjects [Wexler, 2003]. Therefore, the above argument still holds.

but that the computation of its tilt is performed in an allocentric reference frame, contrary to using only retinal data which are egocentric.

On the other hand, the stationarity assumption runs into problems in predicting the effects of shear. At first, all seems well: we take a circle in space with an arbitrary slant and tilt, and rotate it about an arbitrary frontal axis. When we average the square length of 3D displacements generated by this rotation, we find the following expression

$$\frac{r^2}{4} [5 - \cos^2 \sigma - \sin^2 \sigma (2 \cos 2\eta - 1)] \sin^2 \frac{\alpha}{2}. \quad (5)$$

Equation (5) shows that non-stationarity rises with the shear, η , which would seem to be in agreement with our tilt error results in IMMOB (see figure 5). However, our virtual objects were not circles in *space* but in the *image plane*, which were then projected onto the simulated surface; therefore, in space, these objects were ellipses. When we perform the above calculation for these elliptical objects (in parallel projection), we find the following mean square displacement:

$$\frac{r^2}{\cos^2 \sigma} \sin^2 \frac{\alpha}{2}, \quad (6)$$

which is independent of shear. (In perspective projection, the first correction to equation (6) is in second order, which can be safely ignored for our small stimuli.)

Therefore, the stationarity assumption seems to be in agreement with some general features of our data, but not with the dependence of tilt perception on shear.

4.2.3 Optic flow and shear

No argument based solely on optic flow could account for the differences in tilt errors in the ACT and IMMOB, as optic flow was held constant across these conditions. Nevertheless, could an optic flow-based model account for large effect of shear for immobile observers, and for the much smaller but nevertheless significant effect in active observers?

To study this question we used the well-known model of Longuet-Higgins and Prazdny [1980], which assumes a plane undergoing rigid motion, and yields 3D structure and motion from first and second spatial derivatives of optic flow at one point. Here we assume that errors are caused by noise in estimating these derivatives, and further that errors in first derivatives are negligible compared to those in second derivatives. We thus perturb the standard model [Longuet-Higgins and Prazdny, 1980] by the following matrix

$$\begin{pmatrix} 0 & 0 & \mu_1 \\ 0 & 0 & \mu_2 \\ \mu_1 & \mu_2 & 0 \end{pmatrix} \varepsilon \quad (7)$$

and calculate the first-order effect (i.e., keeping only terms up to order ε^1) of this perturbation on the calculated surface normal. We find that the tilt error thus induced is

$$\Delta\tau = \frac{\cos^2 \sigma}{1 - \sin^2 \sigma \cos^2 \eta} \mu_2 \varepsilon. \quad (8)$$

(This perturbation theory result is borne out by numerical simulations with reasonable parameters.) Contrary to our experimental findings, which show a large increase of tilt error with increasing shear angle η in the IMMOB condition, and a smaller but still significant increase in the ACT condition, equation (8) predicts a *decrease* of tilt error with increasing shear. Therefore this well-known SfM model, perturbed in a reasonable way, does not account for the effect of shear on tilt errors, even in the case of the immobile observer.

4.2.4 Linearization of SfM by self-motion information

On a functional level, the problem of simultaneously solving for 3D depth *and* motion of a moving plane from optic flow is a nonlinear one. Indeed, if we take the origin to be at the eye, \mathbf{R} and \mathbf{T} the rotation and translation of the plane, (x, y) a point on the retina, and Z the z -coordinate of the plane in that direction, we have the following optic flow:

$$u_x(x, y) = \frac{1}{Z}(T_x - xT_z) + R_yx^2 - R_xxy - R_zy + R_y \quad (9)$$

$$u_y(x, y) = \frac{1}{Z}(T_y - yT_z) - R_xy^2 + R_yxy + R_zx - R_x. \quad (10)$$

If the flow $u_{x,y}$ is known and the goal is to solve for 3D structure (Z) and motion (\mathbf{R} , \mathbf{T}), equations (9) and (10) are a complex nonlinear system, due to the \mathbf{T}/Z terms.

However, if the motion is known—for instance, if the object is assumed to be stationary in an allocentric frame and self-motion information about \mathbf{R} and \mathbf{T} is integrated into the process—the reduced SfM problem of solving equations (9) and (10) for Z becomes simple and linear. In our experiments, any self-motion information would have to be extra-retinal, as optic flow was the same in the ACT and IMMOB conditions. We hypothesize that quantitative, extra-retinal information about self-motion is integrated into the SfM process.

In our experiments, we have found that in the immobile condition tilt errors had a strong dependence on shear (figure 5). In the active condition, on the other hand, there is a much milder dependence on shear, but a strong dependence on perceived allocentric motion (i.e., in reversal trials, see figure 6). Both of these observations can be taken as indirect evidence for our hypothesis. The strong shear dependence in immobile observers—with performance that approaches chance level for large shear—can be taken as a clue to the complexities of solving a nonlinear problem. According to our hypothesis, the problem is linearized in the active condition, where indeed we find a much reduced dependence of tilt errors on shear. Furthermore, the sharp rise of tilt errors in reversal trials in the active conditions hints that the visual system assumes that relative motion is due to self-motion; that is, that the object is stationary.

4.2.5 Summary of theoretical arguments

In this sub-section we have shown that the allocentric change in the simulated surface normal in the immobile condition, and its constancy in the active condition, can account

neither for the active-immobile differences in tilt error, nor for the different effects of shear on this error in the two self-motion conditions. The stationarity assumption, by itself, does not seem to be able to account for the effects of shear, either. Furthermore, while the active-immobile differences rule out any explanation based solely on optic flow (since optic flow is the same in the two conditions), it might be hoped that traditional models could account for the effect of shear in the immobile condition; however, a perturbed version of the Longuet-Higgins model does not predict the effects of shear in the immobile (or in the active) condition, either.

Finally, although we cannot offer any direct proof, it seems reasonable to assume that the dramatic effect of shear on the perception of surface tilt from optic flow in the *immobile* condition is due to the difficulties attendant upon solving what is intrinsically a nonlinear problem—or evidence that the visual system does not solve this problem with anything like complete generality. The almost complete disappearance of the deterioration of tilt reponse in the *active* condition could be taken as a sign of the linearization of the problem from self-motion information and the assumption of object stationarity.

4.3 Perception and influence of slant

As expected, slant responses are higher with higher simulated slant, showing an overestimation at low slant values and an underestimation of high slant values. Surprisingly, this was also true for the IMMOB condition. The fact that slant is still well recovered in the IMMOB condition—while there are no means by which the observer can scale the visual information with extra-retinal information—indicates that there are some sources of information within the optic flow which help to form an acceptably correct percept.

For example, the observers could have used a heuristic method based on the amount of optic flow, since, in general, the more slanted a plane is, the more optic flow it generates [Todd and Perotti, 1999]. Alternatively, second-order information could have disambiguated the, in first order optic flow, ambiguously defined slant, or the visual system might have used the amount of deformation flow present in the optic flow as a indicator of the amount of slant [Domini and Caudek, 1999].

With our experimental design it not possible to derive which of the former mechanisms is at work. What we do see, however, is that slant perception is better correlated with the simulated slant in ACT than IMMOB, indicating that like with tilt perception, active vision increases the accuracy of slant perception as well.

Simulated slant had a marked influence on the correctness of tilt perception. With an increased slant, subjects generally had smaller errors in tilt estimation, indicating that they were better able to recover the orientation of the stimulus. This fact can be easily explained, because tilt is better defined for higher values of slant. For example, the plane normal is mis-estimated by δ in a random direction, the resulting tilt error is of order $\delta/\sin\sigma$, where σ is the slant.

4.4 Anisotropies in the perception of a plane

4.4.1 A preference for ceiling-like planes

Our results show that two perceptual anisotropies constitute an observer’s preference for ceiling-like planes. Firstly, an anisotropy in reversal rate was found, showing that reversals occurred more often when the normal of the plane was directed upwards (with a maximum at about 75°), reversing a ‘floor’ plane into a ‘ceiling’ plane—see figure 7, right panels. Secondly, an anisotropy in the tilt responses showed that the observers’ perception of tilt was on average more vertical (nearer to 90° or 270°) than the tilt of the simulated plane, with errors up to 18° —see figure 7, left panels.

Reports using other depth cues than optic flow give different results. In stereoscopic slant detection experiments, floor-like surfaces appear to be more slanted than ceiling-like ones [Pierce et al., 1998, Allison et al., 1998], which should make them easier to detect. On the other hand, in a visual search task, using direction of lighting as a depth cue, it has been shown that upward-tilted cubes (comparable to ceiling-like planes) were found faster than downward-tilted ones [Previc and Naegele, 2001]. However, when depth was given by perspective cues [Mamassian and Landy, 1998], observers preferentially saw ground-like planes. Consequently, it seems possible that the type of depth cue determines the bias for floor- or ceiling-like planes.

Anisotropies found in the perception of 2D motion direction detection look similar to the bias towards vertical tilt directions (or horizontal surfaces) in our experiment. Using simple line or random dot stimuli, it has been shown that observers prefer motion directed in one or more of the four cardinal directions [Andrews and Schluppeck, 2000, Rauber and Treue, 1998, Loffler and Orbach, 2001, Bauer and Jordan, 1993]. However, these studies concern 2D motion patterns, and since in our study object orientation was independent from dot motion direction, a causal relationship with the orientation preferences found in our experiment is very unlikely.

Furthermore, it has recently been found that most MSTd cells of the macaque monkey are tuned to report a 90° (floor-like) tilt [Sugihara et al., 2002]. This finding suggests that the significantly greater amount of cells tuned to report vertical tilts embodies the preference for perceiving horizontally orientated planes. Because in this study the axis of rotation and the tilt of the plane were fixed relative to each other, however, a direct extrapolation to our results is not feasible.

Finally, Gibson [1950] argued from an ecological point of view that the visual system is adapted to certain surface orientations because it encounters them more frequently than others. This would then, following Gibson, especially be so for ground surfaces. However, Coppola et al. [1998] has shown that in normal outdoor situations, contours directed in the four cardinal directions (i.e. up, down, left, and right) occur with about equal frequency. Therefore, no bias towards any of these directions should be expected, following Gibson’s reasoning.

Whatever the origin of the preference for ceiling-like planes, our report is, to our knowl-

edge, the first to demonstrate an anisotropy of tilt response in a psychophysical SfM task.^{3,4} This preference can be decomposed into (i) an anisotropy in reversal rate, and (ii) a systematic bias towards vertically orientated tilts.

4.4.2 Motion anisotropy

The finding that the VERT condition generally had greater tilt errors than HORIZ, combined with some subjects' indications that up/down movements were harder to perform, raised the possibility that the complexity of the motion task in the VERT condition caused greater errors. There are two ways that this could have happened. First, if the movement was hard to perform, motion trajectories could differ from HORIZ and therefore the visual stimulus would differ. However, we analyzed movement trajectories and found that their differences were not related to differences in the responses. Second, subjects could have been preoccupied with the motor task in the VERT condition, which could have interfered with the main task, namely indicating the orientation of the stimulus. It seems unlikely that this could be the cause of the VERT-HORIZ effect, as the difference is also present in the IMMOB condition, in which there was no motor task. However, the difference in tilt errors between the two DIRECTION conditions in IMMOB was not significant; nevertheless, it was found that at every individual shear angle the difference between the error in VERT and HORIZ in IMMOB had approximately the same size as the difference in the ACT condition. This suggests that the mechanism causing these differences has nothing to do with the motility of the subject or object, but instead is more likely the result of the visual hardware or a cognitive bias. However, since the DIRECTION effect is present but not significant in the immobile condition, we cannot exclude that the effect was due to interference with the motor task.

5 Summary

This report presents, for the first time, results on the perception of surface orientation during active observer motion. We compare this performance to that in the same subjects experiencing the same optic flow, but while remaining still. The main result is that the error in tilt perception is significantly reduced in the active condition, as compared to the immobile condition. Furthermore, perceived slant is better correlated with simulated slant in the active condition. Since optic flow is the same in the active and immobile conditions,

³Oomes and Dijkstra [2002] found the oblique effect for tilt in their study. However, in their task observers had to align the tilt of the axis of rotation with a predefined direction, and not—like in our task—reproduce the tilt of the object itself. Moreover, calculating the circular standard deviations like the authors did, we found no consistent oblique effect between subjects and between different conditions in our data.

⁴The data of Domini and Caudek [1999] might show a vertical bias as well—or maybe the shear effect—as visual inspection of their figure 12 reveals. However, because of the way they analyzed their data—calculating a linear regression over the entire range of simulated tilt—these authors concluded that their data correlated very well to the simulated tilt.

these results demonstrate the contribution of extra-retinal information concerning self-motion to the perception of 3D structure.

With increasing presence of shear in optic flow, the precision of tilt perception decreased (i.e. the shear effect). However, this effect was severely attenuated in the active relative to the immobile condition. These findings cannot be explained by models based solely on the rigidity assumption, the stationarity assumption or optic flow differences. We speculate that it could be due to the linearization of the structure-from-motion problem by extra-retinal self-motion information, coupled with the assumption of object stationarity.

Furthermore, we have found a systematic bias for perceiving ceiling-like planes. This bias could be decomposed into two anisotropies, one being a non-uniform distribution of tilt reversals, showing a higher rate of reversals when the simulated plane was floor-like, the other being a systematic attraction of responses towards horizontal surfaces.

Previous comparisons between 3D visual perception in active and immobile observers have either involved tasks that are not possible for the latter (such as distance perception), or differences between active and immobile observers due to differing frequencies of choice between two discrete solutions. To our knowledge, this is the first psychophysical result demonstrating a task which can be performed by both active and immobile observers, but which is performed with higher precision in active vision.

A Replay of the active trial

Here we summarize the algorithm used to generate the same optic flow in the immobile condition as in the active condition. The active observer moves both the head and eye (the other eye being covered), in order to fixate the allocentrically stationary fixation point: in other words, the eye undergoes a simultaneous translation and rotation. We will use two reference frames: the allocentrically stationary world frame (whose origin is the fixation point, and which is used by default), and the eye frame, whose origin is the center of the eye, and which rotates and translates with eye. When we say that an active and an immobile trial had the same optic flow, we mean that at every moment (in practice, for every monitor frame) during the two trials, the stimulus was the same in the two corresponding eye frames.

We assume that the eye rotations follow Listing’s law, namely that if the eye is at position $\mathbf{p}_0 = (0, 0, z_0)$ and fixates the origin, and then moves to point $\mathbf{p} = \rho(\sin \theta \cos \phi, \sin \theta \sin \phi, \cos \theta)$ while still fixating the origin, it rotates about an axis parallel to $\mathbf{p}_0 \times \mathbf{p}$. In this ‘law’, rotations about the line of sight are neglected. The corresponding rotation matrix is:

$$\mathcal{L}(\mathbf{p}) = \begin{pmatrix} \cos \theta \cos^2 \phi + \sin^2 \phi & (\cos \theta - 1) \cos \phi \sin \phi & \sin \theta \cos \phi \\ (\cos \theta - 1) \cos \phi \sin \phi & \cos \theta \sin^2 \phi + \cos^2 \phi & \sin \theta \sin \phi \\ -\sin \theta \cos \phi & -\sin \theta \sin \phi & \cos \theta \end{pmatrix} \quad (11)$$

Since \mathcal{L} is orthogonal, $\mathcal{L}^{-1} = \mathcal{L}^T$.

Now, consider that, at a given moment during an active trial, the eye is at point \mathbf{p} , having rotated according to Listing’s law from point \mathbf{p}_0 . If \mathbf{r} is a point on the virtual

stimulus (in the world frame), where is it in the eye’s frame (\mathbf{r}_e)? Since the eye’s frame is parallel to the world frame when the eye is on the z -axis (i.e., the rotation matrix between them is the identity), we have $\mathbf{r}_e = \mathcal{L}(\mathbf{p})(\mathbf{r} - \mathbf{p})$.

At the same moment in the IMMOB trial, the eye is at \mathbf{P} and fixates the origin. Define point \mathbf{R} so that it is at the same position in the eye frame in the immobile trial as point \mathbf{r} was in the active trial. In other words, the changes in the active condition must be the same as in the IMMOB condition, in the eye’s frame:

$$\mathbf{r}_e = \mathcal{L}(\mathbf{p})(\mathbf{r} - \mathbf{p}) = \mathcal{L}(\mathbf{P})(\mathbf{R} - \mathbf{P}) \quad (12)$$

Equation (12) guarantees the same optic flow in immobile and active trials. Ideally, the observer in the immobile condition should not move, in which case $\mathcal{L}(\mathbf{P})$ would be the identity, but since we wanted to correct for any spurious motion in the immobile condition, we need a rotation matrix here too. Equation (12) can be easily solved for \mathbf{R} , giving

$$\mathbf{R} = \mathbf{P} + \mathcal{L}^T(\mathbf{P})\mathcal{L}(\mathbf{p})(\mathbf{r} - \mathbf{p}). \quad (13)$$

In deriving equation 13, we have made two assumptions which may be false. First, Listing’s law is not exact. It is off by about 10%, since the eye does rotate about the line of sight. As a consequence, the image does not rotate in the immobile condition while it does so slightly in the active condition. However, rotations in the image plane do not contain 3D information and are as such uninformative for our subjects. The second possible problem would be if subjects did not fixate the fixation point perfectly. This is not accounted for by the rotation matrix, but any discrepancies would only result in rotations about the center of the eye (i.e., wholesale shifts of the retinal image), which again are uninformative about stimulus structure.

References

- R.S. Allison, I.P. Howard, B.J. Rogers, and H. Bridge. Temporal aspects of slant and inclination perception. *Perception*, 27(11):1287–1304, 1998.
- T.J. Andrews and D. Schluppeck. Ambiguity in the perception of moving stimuli is resolved in favour of the cardinal axes. *Vision Research*, 40(25):3485–3493, 2000.
- E. Batschelet. *Circular Statistics in Biology*. Academic Press, London, 1981.
- R. Bauer and W. Jordan. Different anisotropies for texture and grating stimuli in the visual map of cat striate cortex. *Vision Research*, 33(11):1447–1450, 1993.
- M.L. Braunstein. Motion and texture as sources of slant information. *Journal of Experimental Psychology*, 78(2):247–253, 1968.
- D.M. Coppola, H.R. Purves, A.N. McCoy, and D. Purves. The direction of orientated contours in the real world. *Proc. Natl. Acad. Sci. USA*, 95:4002–4006, 1998.

- V. Cornilleau-Pérès, M. Wexler, J. Droulez, E. Marin, C. Miège, and B. Bourdoncle. Visual perception of planar orientation: dominance of static depth cues over motion cues. *Vision Research*, 42(11), 2002.
- T.M. Dijkstra, V. Cornilleau-Pérès, C.C. Gielen, and J. Droulez. Perception of three-dimensional shape from ego- and object-motion: comparison between small- and large-field stimuli. *Vision Research*, 35(4):453–462, 1995.
- F. Domini and C. Caudek. Perceiving surface slant from deformation of optic flow. *Journal of Experimental Psychology: Human Perception and Performance*, 25(2):426–444, 1999.
- J.J. Gibson. *The perception of the visual world*. Houghton-Mifflin, Boston, 1950.
- D.D. Hoffman. Inferring local surface orientation from motion fields. *Journal of the Optical Society of America*, 72:888–892, 1982.
- G. Loffler and H.S. Orbach. Anisotropy in judging the absolute direction of motion. *Vision Research*, 41(27):3677–3692, 2001.
- H.C. Longuet-Higgins and K. Prazdny. The interpretation of a moving retinal image. *Proceedings of the Royal Society of London (B, Biological Sciences)*, 208(1173):385–397, 1980.
- P. Mamassian and M.S. Landy. Observer biases in the 3D interpretation of line drawings. *Vision Research*, 38(18):2817–2832, 1998.
- T.S. Meese, M.G. Harris, and T.C.M. Freeman. Speed gradients and the perception of surface slant: analysis is two-dimensional not one-dimensional. *Vision Research*, 35(20):2879–2888, 1995.
- J.F. Norman, J.T. Todd, and F. Phillips. The perception of surface orientation from multiple sources of optical information. *Perception and Psychophysics*, 57(5):629–636, 1995.
- H. Ono and M.J. Steinbach. Monocular stereopsis with and without head movement. *Perception and Psychophysics*, 48(2):179–187, 1990.
- A.H.J. Oomes and T.M.H. Dijkstra. Object pose: perceiving 3-D shape as sticks and slabs. *Perception and Psychophysics*, 64(4):507–520, 2002.
- F. Panerai, V. Cornilleau-Pérès, and J. Droulez. Contribution of extraretinal signals to the scaling of object distance during self-motion. *Perception and Psychophysics*, 64(5):717–731, 2002.
- F. Panerai, S. Hanneton, J. Droulez, and V. Cornilleau-Pérès. 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, 1999.

- C.-H. Peh, F. Panerai, J. Droulez, V. Cornilleau-Pérès, and L.-F. Cheong. Absolute distance perception during in-depth head movement: Calibrating optic flow with extra-retinal information. *Vision Research*, 42(16):1991–2003, 2002.
- B.J. Pierce, I.P. Howard, and C. Feresin. Depth interactions between inclined and slanted surfaces in vertical and horizontal orientations. *Perception*, 27(1):87–103, 1998.
- F.H. Previc and P.D. Naegle. Target-tilt and vertical-hemifield asymmetries in free-scan search for 3-D targets. *Perception and Psychophysics*, 63(3):445–457, 2001.
- H.J. Rauber and S. Treue. Reference repulsion when judging the direction of visual motion. *Perception*, 27(4):393–402, 1998.
- B. Rogers and M. Graham. Motion parallax as an independent cue for depth perception. *Perception*, 8(2):125–134, 1979.
- S. Rogers and B.J. Rogers. Visual and nonvisual information disambiguate surfaces specified by motion parallax. *Perception and Psychophysics*, 52(4):446–452, 1992.
- K.A. Stevens. Surface tilt (the direction of slant): a neglected psychophysical variable. *Perception and Psychophysics*, 33(3):241–250, 1983.
- H. Sugihara, I. Murakami, K.V. Shenoy, R.A. Andersen, and H. Komatsu. Response of MSTd neurons to simulated 3D orientation of rotating planes. *Journal of Neurophysiology*, 87(1):273–285, 2002.
- J.T. Todd and V.J. Perotti. The visual perception of surface orientation from optical motion. *Perception and Psychophysics*, 61(8):1577–1589, 1999.
- S. Ullman. *The interpretation of visual motion*. MIT Press, Cambridge, Mass., 1979.
- W.J. van Damme and W.A. van de Grind. Non-visual information in structure-from-motion. *Vision Research*, 36(19):3119–3127, 1996.
- H. von Helmholtz. *Handbuch der Physiologischen Optik*. Voss, Hamburg, 1867.
- H. Wallach and D.N. O’Connell. The kinetic depth effect. *Journal of Experimental Psychology*, 45:205–217, 1953.
- H. Wallach, L. Stanton, and D. Becker. The compensation for movement-produced changes in object orientation. *Perception and Psychophysics*, 15:339–343, 1974.
- M. Wexler. Allocentric perception of space and voluntary head movement. *Psychological Science*, 2003. In Press.
- M. Wexler, I. Lamouret, and J. Droulez. The stationarity hypothesis: an allocentric criterion in visual perception. *Vision Research*, 41(23):3023–3037, 2001a.
- M. Wexler, F. Panerai, I. Lamouret, and J. Droulez. Self-motion and the perception of stationary objects. *Nature*, 409:85–88, 2001b.