Available online at www.sciencedirect.com
SCIENCE DIRECT
Vision Research 46 (2006) 3494–3513
Vision Research
www.elsevier.com/locate/visres

The visual perception of plane tilt from motion in small field and large field: Psychophysics and theory

H. Zhong [a], V. Cornilleau-Pérès [b,c,*], L.-F. Cheong [d], G.M. Yeow [d], J. Droulez [e]

[a] Department of Cognitive Sciences, University of California, Irvine, USA
[b] Image Processing and Application Laboratory, CNRS-NUS, Singapore
[c] Labo. Neurosciences Fonctionnelles and Pathologies UMR8160, CNRS-Univ.Lille2-CHRU-Lille, France
[d] Department of Electrical and Computer Engineering, National University of Singapore, Singapore
[e] Labo. de Physiologie de la Perception and de l'Action, CNRS-Collège de France, Paris, France

Received 18 November 2005; received in revised form 30 March 2006

Abstract

Subjects indicated the tilt of dotted planes rotating in depth, in monocular viewing, under perspective projection. The responses depended on the FOV (field of view) and on the angle W between the tilt and frontal translation (orthogonal to the rotation axis). Response accuracy increased with the FOV, and decreased with W. Our results support the processing of the second-order optic flow in all cases, but indicate that this flow is quantitatively small in small-field, leading to tilt ambiguities. We examine computational models based on the affine components of the optic flow to interpret our results.

Keywords: Vision; Motion parallax; Plane; Depth; Orientation; 3D space; Optic flow; Tilt Direction; Structure from motion

1. Introduction

Motion parallax is a visual depth cue that contributes to the perception of the 3D space around us. For instance, it has been shown that monocularly enucleated subjects tend to rely on head movements to compensate for their lack of binocular vision (Marotta, Perrot, Nicolle, Servos, & Goodale, 1995), and that specific cortical circuits are dedicated to the processing of motion variations in the visual scene (reviews in Cornilleau-Pérès & Gielen, 1996, or Lappe, 2000). In parallel, many computer vision studies have addressed the problem of computing 3D structure from motion parallax in image sequences (Longuet-Higgins & Prazdny, 1980; Waxman & Ullman, 1985).

The visual perception of surface orientation is required for navigation, when climbing a slope for instance, or for actions like grasping (Dijkerman, Milner, & Carey, 1996). The orientation of a plane relative to the eye can be described by the slant and tilt (Stevens, 1983). Calling **N** the vector normal to a plane (Fig. 1), the slant σ quantifies the global slope of the surface, relative to the line of sight, and corresponds to the angle between **N** and this line. The tilt τ is the direction of maximal slope, i.e., the direction of the projection of **N** in the frontoparallel plane (Figs. 2A and B). Tilt computation requires only the computation of ordinal relationships between object points (Garding, Porrill, Mayhew, & Frisby, 1995; Koenderink & van Doorn, 1995). On the opposite, the calculus of slant requests a metrical representation of depth in the human visual system. Note that this metrical representation of depth should not be confounded with absolute distance perception, which is normally based on the estimation of distances scaled from the subject's own body for instance (e.g., the perception of absolute distances from self-motion, or binocular disparity are scaled by quantities such as the body displacements, the convergence angle, and interocular distance).

* Corresponding author. Tel.: +33 6 61 38 83 79.
E-mail address: c.peres2@wanadoo.fr (V. Cornilleau-Pérès).

0042-6989/$ - see front matter © 2006 Elsevier Ltd. All rights reserved.
doi:10.1016/j.visres.2006.04.003

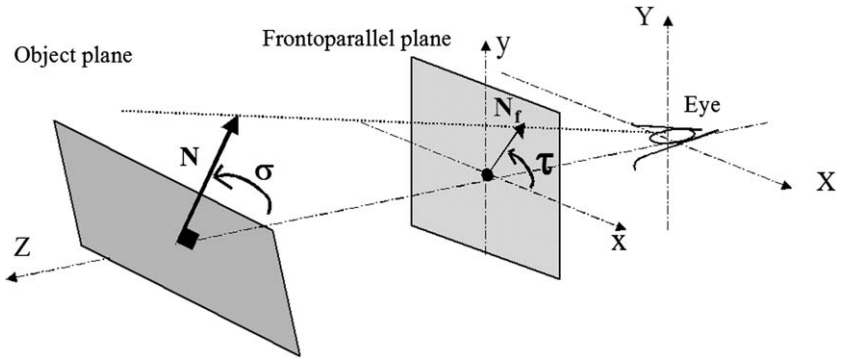

Fig. 1. A plane of normal **N** is seen in perspective projection, with the perspective centre located at the observer's eye. Z is the axis from the eye to the plane point, where the plane orientation is to be analysed. With Z, the axes X and Y form a coordinate system (XYZ) of the 3D space, and project as axes x and y onto the frontoparallel plane, perpendicular to Z. The vector $\mathbf{N_f}$ is the projection of **N** onto the (xy) plane. $\mathbf{N_f}$ makes an angle τ (the tilt) with the axis x. **N** forms an angle σ (the slant) with the axis Z.

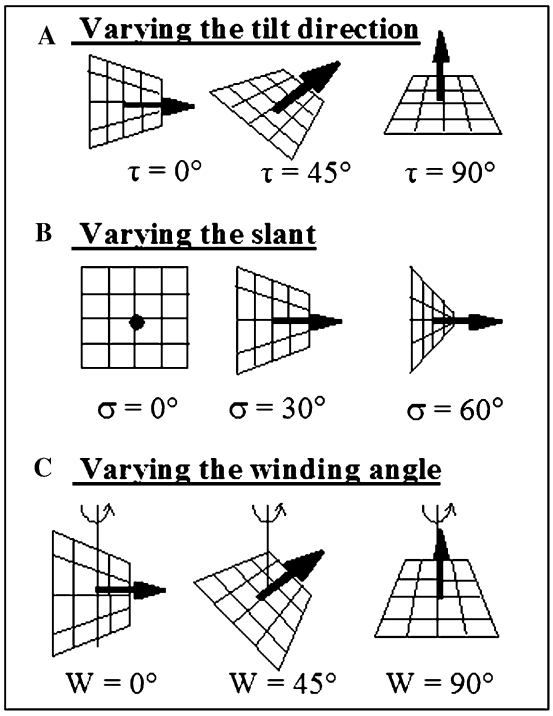

Fig. 2. Illustration of the variations of different variables. (A) Varying the tilt, i.e., the direction of maximal depth gradient. (B) Varying the slant, i.e., the angle between the object plane and the frontoparallel plane. (C) Varying the motion/orientation configuration. The winding angle W is the angle between the tilt direction, and the component of frontoparallel translation (orthogonal to the rotation axis).

The human visual perception of slant has been intensively explored (e.g. Beck & Gibson, 1955; Börjesson & Lind, 1996; Braunstein, 1968; Freeman, 1966; Freeman, Harris, & Meese, 1996; Gibson & Carel, 1952). By comparison, and in spite of the growing importance attributed to the perception of ordinal depth (Koenderink & van Doorn, 1995), the psychophysical studies on tilt perception remain sparser. For surfaces with multiple orientations, Norman, Todd, and Phillips (1995) found a high correlation between the stimulus and perceived tilt, when 3D shape is specified by motion, stereopsis or texture. For motion parallax, Domini and Caudek (1999) and Todd and Perotti (1999) found that observers estimate tilt more accurately than slant in multiple-view stimuli. However, they used a single direction of 3D motion (a rotation about a vertical axis), which may bias the results (any information about the direction of the 3D motion may simplify the resolution of the 3D structure from motion problem). They did not explore the general case where both the motion and tilt direction vary randomly across trials. They also used an orthographic projection, which is an approximation of perspective projection, valid only for small viewing angles.

Previously, we observed that reports of plane tilt from motion parallax depend on the orientation of the tilt relative to the motion (Cornilleau-Pérès et al., 2002). Consider a rotation in depth about a frontoparallel axis. Such a motion involves a component of frontal translation $\mathbf{T_f}$ orthogonal to the rotation axis (intuitively, it indicates the overall motion direction projected in a frontoparallel plane). We defined the winding angle W, as the angle between the tilt and $\mathbf{T_f}$ (Fig. 2C), and found that the error on the perceived tilt strongly increases with W in small field (8° width), but more weakly so in large field (60° width). This result has been confirmed by others (van Boxtel, Wexler, & Droulez, 2003), but has not received any computational account, so far.

Yet, models of tilt computation from motion have been proposed (Koenderink & van Doorn, 1976; Longuet-Higgins, 1984; Subbarao, 1988). They indicate the existence of multiple solutions (generally two) in the problem of tilt computation from optic flow. These approaches express the optic flow (retinal image velocity) projected by a moving plane as a second-order polynomial vector of the image coordinates.

The present study aims at studying in details the dependence of tilt perception from the winding angle W, in small and large field, and at exploring the computational correlates of this property. We start by recording the perceived orientation of a moving plane in small (8°) and large (60°) field of views. Because multiple-frame stimuli may provide

complementary acceleration information (Hoffman, 1982), we use two-view stimuli. We then study the exact shape of the tilt response distributions for different W values. On the theoretical side, we show that the second-order terms of the flow are small, relative to the first order terms in our small-field stimuli. We analyze the 3D content information contained in the first-order terms, or affine flow. We describe the associated 3D ambiguities, and study different constraints that could be associated by the visual system to the affine flow to lift these ambiguities. The corresponding models predict different shapes of the tilt response distributions in small and large field. The confrontation of these predictions with our results leads to support the existence of a processing of the full flow, including second-order terms, in our large field conditions. In small-field, the weakness of the second-order terms seems to induce an analysis of the first-order (affine) flow, possibly associated with the stationarity hypothesis proposed by others (Wexler, Panerai, Lamouret, & Droulez, 2001).

2. Geometrical preliminaries

The position of the eye is the origin of a (XYZ) coordinate system, with the Z-axis lying along the line of sight (Fig. 1). (xy) is the axis system of the frontal plane, projected from the system (XY). Let P be an object plane of equation

$$Z = Z_X \cdot X + Z_Y \cdot Y + Z_0. \quad (1)$$

We exclude the case where P is parallel to the Z-axis, because in that case it projects as a single line onto the frontal plane. $\mathbf{N}$ is the normal to P, pointing toward the eye, with coordinates $(Z_X, Z_Y, -1)$. Its projection in the frontal plane is the vector $\boldsymbol{\tau}$ of coordinates (Z_X, Z_Y) in (xy).

The angle between $\boldsymbol{\tau}$ and the x-axis is the tilt angle τ (Fig. 2A). The angle between $\mathbf{N}$ and the Z-axis is the *slant angle* σ (Fig. 2B). Under the axis orientation specified in Fig. 1, calling $s = \mathrm{tg}\,\sigma$, we have:

$$\begin{aligned} Z_X &= s \cdot \cos\tau, \\ Z_Y &= s \cdot \sin\tau \end{aligned} \quad (2)$$

The 3D motion of P can be decomposed in a rotation $\boldsymbol{\Omega}$ around the eye, of coordinates $(\Omega_X, \Omega_Y, \Omega_Z)$, and a translation $\mathbf{T} = (T_X, T_Y, T_Z)$. The vector $\mathbf{T_f} = (T_X, T_Y)$ represents the frontal translation. Note that if the plane rotates around a frontoparallel axis, $\mathbf{T_f}$ is orthogonal to that rotation axis (Cornilleau-Pérès & Droulez, 1989). The angle between $\mathbf{T_f}$ and $\boldsymbol{\tau}$ is called the winding angle W (unsigned, ranging between 0 and 90°, see Fig. 2C).

3. Psychophysics

3.1. Experiment 1

Subjects reported the tilt of a plane presented in 2 frame image sequences in SF (small-field, visual angle 8°) and LF (large-field, visual angle 60°) in monocular vision. The orientation of the plane was indicated by the adjustment of a graphical probe, superimposed on the moving stimulus, a method inspired from Stevens (1983). Motion parallax was generated through a rotation in depth of the plane, about a frontoparallel axis of random direction. This motion was chosen because

(1) It contains a component of frontoparallel translation T_f proportional to the tangent of the 3D rotation angle, and a component of rotation about the observer's eye (the latter being devoid of any depth information). T_f is known to generate a reliable motion parallax to the visual system, from a theoretical point of view. In SF, for a static observer, the detection of unsigned depth gradients is more accurate for a rotation in depth (T_f embedded in a rotation in depth), than for a pure frontal translation T_f (Dijkstra, Cornilleau-Pérès, Gielen, & Droulez, 1995). In LF, both 3D movements (a rotation in depth involving T_f, or a pure T_f) lead to similar sensitivities to motion parallax. Therefore, the rotation in depth represents an optimal condition for the detection of motion parallax.
(2) It maintains the central point of the stimulus static on the screen, which is convenient to adjust a graphical probe and indicate the surface orientation.

3.1.1. Methods

3.1.1.1. Subjects. Nine observers aged 21–28 served as naïve subjects. All of them had normal or corrected-to-normal vision, and gave their written informed consent as to the goal of the experiments.

3.1.1.2. Design. We examined the effects of two variables on the judgments of tilt and slant: (1) the size of the visual stimulus (diameter 8° or 60° visual angle) and (2) the winding angle W randomised between 0° and 90°. Within the frontoparallel plane, the tilt τ and the angle of the rotation axis randomly varied between 0° and 360°. Therefore, the winding angle W defined above varied randomly between 0° and 90° (the frontoparallel translation oscillates back and forth during the trial and thus one cannot separate out responses due to a positive or negative sign alone). The slant σ of the plane was 35°, and the rotation amplitude between the two views was 3°. There were 8 sessions of 108 trials for each field size. The SF and LF sessions were performed alternately in random order.

3.1.2. Apparatus

The stimulus patterns were generated on a PC computer, and displayed either on the 19-in. monitor (SF), or on a rigid glass-fabric screen, using an Electrohome Marquee Ultra8500 projector (LF). The stimuli had a diameter of 728 pixels. We used an anti-aliasing software to achieve subpixel accuracy, each dot covering a 3×3 pixels area. The refresh rate was 85 Hz.

3.1.3. Stimuli

The viewing distance was 1.96 m (SF), and 1.73 m (LF), with stimulus diameters of 27.5 cm and 2 m, respectively. The stimuli were perspective projections of dotted planes (Fig. 3A), with the centre of the perspective projection located at the subject's eye E. The centre of the stimulus, on the screen, called point K, was located at the subject's eye level. The line EK intersected the stimulus plane in K. The stimulus plane, defined by its tilt and slant, rotated about a frontoparallel axis, passing through point K. In this configuration, the whole stimulus projection is scale invariant, and does not depend on the simulated viewing distance.

The two views of the plane were calculated so that the dot distribution had a uniform density *on the projection plane* (i.e., in the display screen, Fig. 3A) for the *median position* of the surface, defined as the intermediate position between the 2 views. Also, this median position corresponded exactly to the tilt τ and slant σ chosen for the stimulus. Hence, the 2 views corresponded to tilt and slant values which could differ by a maximum of 1.5° from the median τ and σ, because they corresponded to a rotation of $-1.5°$ and 1.5° of the plane, from its median position. The duration of each view was 0.38 ± 0.015 s. The dot number was 572 ± 17. The mean dot speed was 0.03°/s in SF, and 0.66°/s in LF.

We chose to use the same total resolution in SF and LF, i.e., an angular resolution that scaled with the stimulus size, and the same dot number, so as to maintain a similar amount of discretized visual information across the SF and LF conditions. This was initially due to technological constraints, because we could not increase the resolution in LF. It turned out to be the right choice, because if we had used the central low resolution of the projection screen in SF, the lower performance observed in SF, as compared to LF, could have been attributed to this low resolution (recall that acuity is maximal in central vision for both position and motion).

The graphic probe adjusted by the subject to indicate the perceived plane orientation consisted of a needle and an ellipse. Subjects adjusted the needle orientation, and the ellipse width with the computer mouse, to indicate the perceived tilt (the direction of the needle) and slant (width of the ellipse) (Fig. 3B), according to Stevens' method. The needle had a maximum length of 194 pixels, and the larger width of the ellipse was fixed at 194 pixels. The mean luminance was 0.23 cd/m^2.

3.1.4. Procedure

The subject was seated in darkness, with head maintained in a chinrest, and an eye patch covering the non-dominant eye. He/she was asked to fixate the centre of the stimulus. The stimulus plane was presented in continuous oscillation (with the 2 views displayed alternatively). After 3 s of presentation, the subject could adjust the position of the mouse to modify the orientation of the probe superimposed on the stimulus. Upon completion of the adjustment, he/she clicked on the mouse, and proceeded to the next trial. The trial duration usually ranged around 8 s.

3.1.5. Data analysis

We partitioned the winding angles in nine intervals: 0°–10°, 10°–20°, ... 80°–90°. The average number of trials for each subject in each W interval was 96 (standard deviation 7). Tilt sign is ambiguously perceived if subjects tend to respond according to the true tilt τ, as well as to the opposite tilt $\tau + 180°$. We measured this ambiguity (tilt reversal) by calculating the percentage of trials where the unsigned tilt error ranges between 90° and 180°. We corrected the responses for the 180° ambiguity (tilt reversal), i.e., if τ_s is the stimulus tilt and τ_p is the reported tilt, we considered the response as either τ_p or $\tau_p + 180°$, whichever is closer to τ_s. Thus, we obtained a "corrected absolute tilt error" ranging between 0° and 90°, as a measure of the performance.

For the graphical presentation of the results, we consider the bisector **B** of the vectors $\mathbf{T_f}$ and $\boldsymbol{\tau_s}$ as the origin of the angles. In this convention, if $\mathbf{T_f}$ is at angle W from $\boldsymbol{\tau_s}$, the angle $\boldsymbol{\tau_s}$ is set to $-W/2$, while $\mathbf{T_f}$ is at $+W/2$. In the case where $\mathbf{T_f}$ is at $-W$ from $\boldsymbol{\tau_s}$, we apply a reflection about **B** to the 3 vectors $\mathbf{T_f}$, $\boldsymbol{\tau_s}$ and $\boldsymbol{\tau_p}$, so that we again affix an orientation $-W/2$ to $\boldsymbol{\tau_s}$ and $+W/2$ to $\mathbf{T_f}$. Hence, for a given value of W, we can group all responses into a single histogram, with $\boldsymbol{\tau_s}$ at $-W/2$ and $\mathbf{T_f}$ at $W/2$.

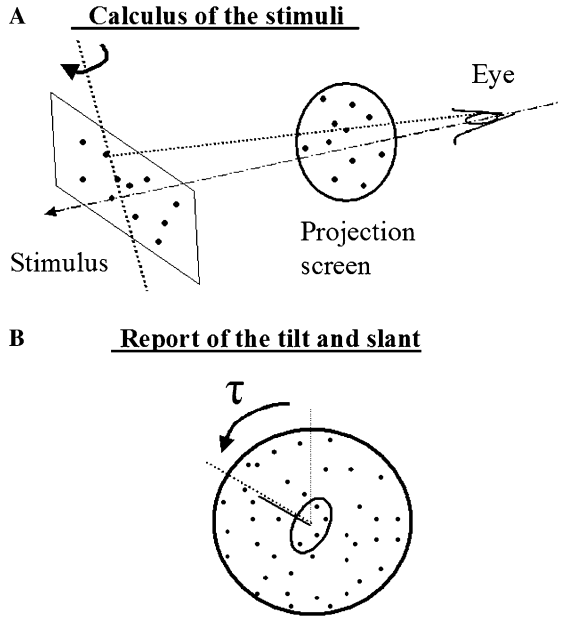

Fig. 3. (A) A dotted plane of a given orientation in space oscillates about a frontoparallel axis between 2 extreme positions. For each of these positions, the projections of the dots onto the frontoparallel screen plane are calculated, using the subject's eye as the centre of perspective projection. (B) The subject sees the projection of the dots located within a circular window. He/she adjusts a graphic probes superimposed on the dot distribution, so as to indicate the perceived orientation of the plane.

Due to the periodicity of the tilt, we used circular statistics to find the mean τ_{pm} of the tilt distributions. For each value τ_0 in [0, 180°] we calculated the mean of the reported tilt within the single period $[\tau_0, \tau_0 + 180°]$. The mean corresponding to the minimum variance was kept as the final mean of the tilt distribution.

As the shape of most of the bisector-oriented distributions is found sharper than the normal distribution, we fitted them with Laplace functions. In some cases, they were also approximated by the sum of two Laplace distributions.

Since most distributions are not strictly normal (for instance the absolute tilt error is bounded by the value 0), we use non-parametric tests, except if otherwise stated. For instance we calculate the Spearman correlation, and we compare independent samples with the Mann–Whitney U test (MWU). The level of significance is taken at $p < 0.05$.

3.1.6. Results

3.1.6.1. Verbal reports. All subjects found the task more difficult in SF than in LF. Eight of the 9 subjects reported a perception of curved surfaces, rather than planes, for large values of W, particularly in LF. The perceived curvature was generally convex (i.e., the plane looks like a bump seen from above).

3.1.6.2. Effect of the field of view (FOV) on the reported tilt sign. We find 41% of tilt reversals in SF and only 2.4% in LF. This confirms previous results showing that perspective projection lifts the 180° tilt ambiguity in LF (Cornilleau-Pérès, Wong, Cheong, & Droulez, 2000; Cornilleau-Pérès et al., 2002) and, more generally, the ambiguity on the depth sign (Dijkstra et al., 1995). All subsequent results are corrected with respect to the tilt sign.

3.1.6.3. Effect of the FOV and W on the absolute tilt error. The average absolute tilt error (Fig. 4B) is always lower in LF than in SF, especially for large values of W. This effect of the FOV is significant for 8 of the 9 subjects (MWU test at $p < 0.01$), and for the whole population (MWU test $Z = 36$, $p < 0.001$).

In SF, the average absolute tilt error increases dramatically with W. This effect is weaker in LF. The Spearman correlation of the absolute tilt error with W is significant for each subject in SF (overall coefficient: 0.472, $p <$ 1.E-6) and for 8 of the 9 subjects in LF (overall coefficient: 0.115, $p <$ 1.E-6).

3.1.6.4. Effect of the FOV and W on the mean reported tilt. A tilt-error can be due either to a lack of accuracy (a shift of the mean reported tilt, away from the stimulus tilt), or to a lack of precision (flattening of the response distribution). Fig. 5 plots the histograms of bisector-oriented tilt reports for each W interval, superimposed with a one-peak Laplace fitting. The vertical dotted lines indicate the abscissae $-W/2$ (stimulus tilt τ_s), and $W/2$ (frontal translation $\mathbf{T_f}$). The origin of the abscissae is the bisector **B**. The mean reported tilt τ_{pm} tends to deviate from the stimulus tilt toward **B**. This effect is large in SF, and weak in LF.

Fig. 6 plots the mean individual perceived tilt τ_{pm} in SF and LF for each subject and W category. The ascending straight line indicates the frontal translation $\mathbf{T_f}$, while the descending line indicates the stimulus tilt τ_s. As described above, τ_{pm} lies between $\mathbf{T_f}$ and τ_s in small field, but with strong variations among subjects. For instance subject

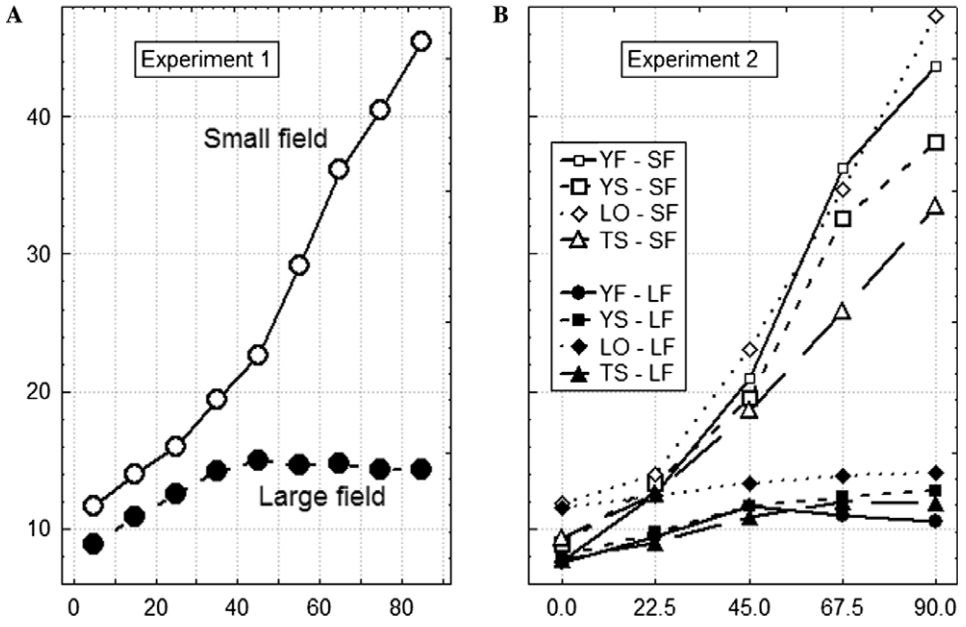

Fig. 4. Experiment 1. Average absolute tilt error in degrees (corrected for the 180° tilt ambiguity) as a function of the angle W (in degrees) in small field (upper curves) and large field (lower curves). (A) For Experiment 1, the data is averaged over all subjects. (B) For Experiment 2, the individual curves are presented. They are averaged for all stimulus velocities and slants.

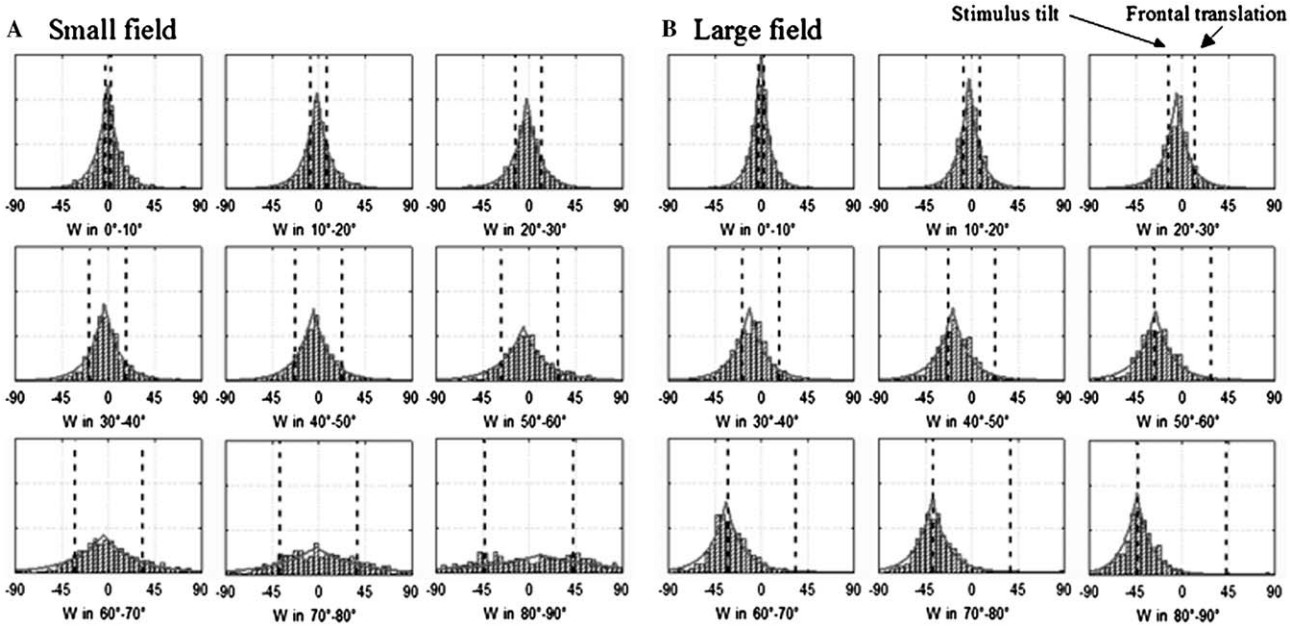

Fig. 5. Experiment 1. Histograms of the tilt responses (corrected for the 180° ambiguity) in Experiment 1, and their Laplacian fitting. The vertical lines indicate the stimulus tilt (left) and the direction of frontal translation (right). Each graph corresponds to a value of the angle W. (A) Small field. (B) Large field.

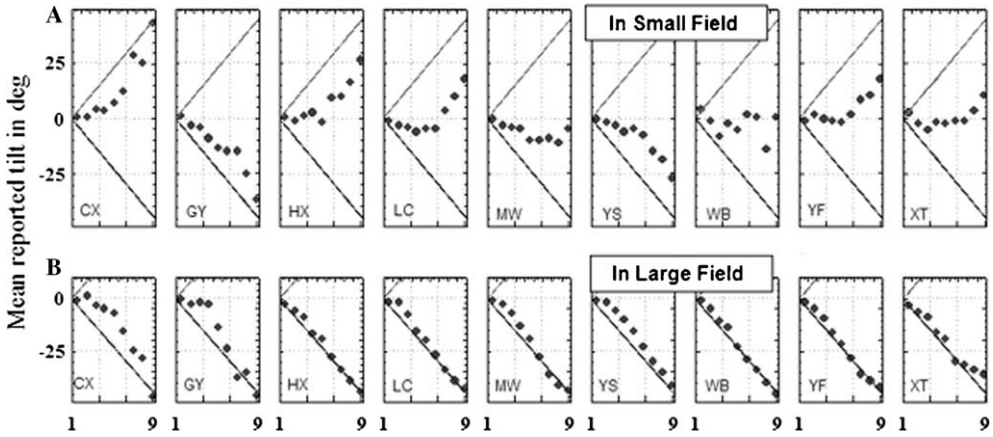

Fig. 6. Experiment 1. Average individual perceived tilt in Experiment 1 (corrected for the 180° ambiguity), as a function of the winding angle (abscissa 1 corresponds to W in [0°–10°], 2 is for [10°–20°], etc.) in small field (upper graphs) and large field (lower graphs). The initials of the corresponding subjects are indicated on each graph.

CX tends to answer according to $\mathbf{T_f}$, GY according to the stimulus tilt τ_s and XT according to **B**. By contrast, in large field, τ_{pm} is close to τ_s for all subjects.

Therefore, τ_{pm} is close to **B** (slightly toward τ_s) for the grouped responses, with a large intersubject variability (the individual τ_{pm} lying between $\mathbf{T_f}$ and τ_s) in SF. In LF, τ_{pm} is close to τ_s. The accuracy of the responses decreases as W increases in SF, but not in LF.

3.1.6.5. Effect of the FOV and W on the tilt response distributions. The response distributions tend to flatten as W increases (Fig. 5), indicating a corresponding decrease in the response precision. This effect is strong in SF, and slight in LF. In the latter case, the distributions present a single peak (Fig. 5B). In small field, we usually observe a one-peak distribution (Fig. 5A) except for $W = 80°$-90°, where the distribution tends to present 2 peaks at $-W/2$ and $+W/2$. Using a two-peak fitting for $80° < W < 90°$ in SF, we find two symmetrical shallow peaks centred at $\mathbf{T_f}$ and τ_s (Fig. 7B) for the grouped responses. The individual response histograms (Fig. 7A) indicate that the general shape of the distributions varies among subjects. Two subjects respond according to τ_s, two others according to $\mathbf{T_f}$, and 4 others present 2 response peaks located at less than 5° from $\mathbf{T_f}$ and τ_s. Therefore, the existence of two equivalent peaks at$\mathbf{T_f}$ and τ_s is not confirmed at the individual level.

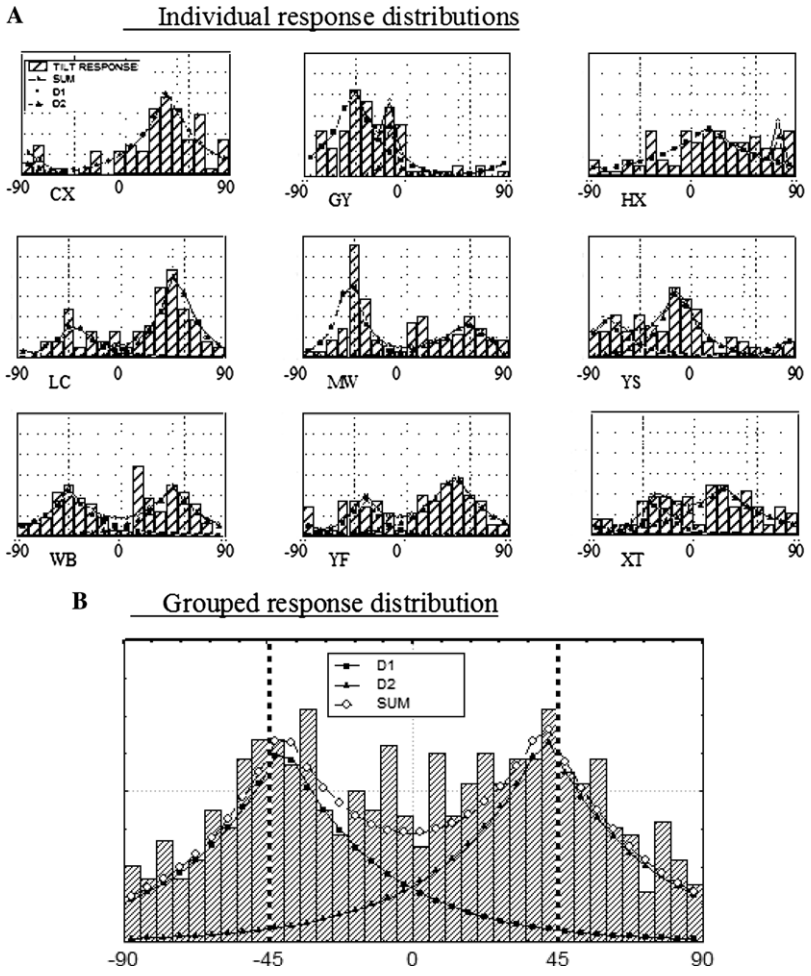

Fig. 7. Experiment1. Response distributions for W in [80°–90°] (corrected for the 180° ambiguity), and their two-peaks Laplacian fitting. D1 and D2 are the 2 Laplacian functions, and SUM is their summed distribution. (A) Individual distributions. (B) Grouped distribution.

In summary, we observe different response patterns in SF and LF. In SF, the increase of the absolute tilt error with W is due to a decrease in accuracy and precision of the responses, with the mean perceived tilt lying close to the bisector of $\mathbf{T_f}$ and $\boldsymbol{\tau_s}$. In LF, the mean reported tilt is close to the stimulus tilt, but the response precision decreases as W increases. The response distributions tend to be one-peak shaped, except when $80° < W < 90°$ in SF where the grouped response distribution presents 2 peaks.

3.2. Experiment 2

In the first experiment, each W category is a 10° interval, instead of a single W value. The second experiment examines the shape of the response distributions for single W values, and tests the effect of plane slant and dot speed on those distributions. Here 'dot speed' refers to the average 2D speed of the plane dots on the display screen.

3.2.1. Methods

They were the same as in Experiment 1, except for the following:

3.2.1.1. Participants. Four observers aged 22–23 served as naïve subjects for this experiment. Two of them had participated in the previous experiment. All had normal or corrected-to-normal vision.

3.2.1.2. Design. We examined the effects of 4 parameters on the judgements of plane orientation in terms of tilt and slant: (1) size of the visual stimulus (diameter 8° or 60° visual angle) (2) W, with 5 values at 0°, 22.5°, 45°, 67.5°, and 90°. (3) plane slant, with 8 values between 17.5° and 35°, with a step of 2.5°. (4) average dot speed V, with 6 speed values for each slant value. There were 10 possible tilt values, which were presented for each set of parameters, making a total of 4800 trials for each subject.

For a constant 3D rotational angle Ra, V increases with slant, as V is roughly the product of the tangents of Ra, and of the slant (Annex B) in SF. In order to vary slant and V independently, we chose a different range of 3D rotation angles for each slant values in small field (Table 1), so as to obtain similar V values across all slant values. The resulting correlation between

Table 1
Values of the slants (left column, in bold), rotation angles (in italics) and mean dot speed (2 bottom rows, in °/s) in Experiment 2

Slant (°)	3D rotation angle (°)					
17.5	*2.22*	*4.44*	*6.65*	*8.86*	*11.07*	*13.28*
20.0	*1.92*	*3.85*	*5.77*	*7.68*	*9.60*	*11.52*
22.5	*1.69*	*3.38*	*5.07*	*6.75*	*8.44*	*10.13*
25.0	*1.50*	*3.00*	*4.50*	*6.00*	*7.50*	*9.00*
27.5	*1.35*	*2.69*	*4.03*	*5.37*	*6.72*	*8.07*
30.0	*1.21*	*2.43*	*3.64*	*4.85*	*6.06*	*7.27*
32.5	*1.10*	*2.20*	*3.30*	*4.40*	*5.50*	*6.60*
35.0	*1.00*	*2.00*	*3.00*	*4.00*	*5.00*	*6.00*
Dot speed LF (± 15%)	0.24	0.47	0.66	0.87	1.07	1.28
Dot speed SF (± 4%)	0.011	0.021	0.031	0.042	0.052	0.062

The 'LF' and 'SF' bottom row gives the mean dot speed over the image in large field and small field, respectively, with the possible variation range (in percentage) in parentheses.

V and slant was not significant (Spearman −0.01). In LF, we kept the same set of rotation angles, and found that V was then significantly correlated with slant, but negatively (Spearman −0.21). Hence, in all conditions, the chosen 3D rotation angles allow to cancel out the positive correlation between slant and V. For each slant and rotation angle, the V values are indicated in the bottom row of Table 1.

In LF, the values of the slant and 3D rotation amplitude were those used in SF. Therefore, the restriction of the calculus of V to the central disk of 8° diameter (i.e., over an area covering the small field stimulus) is identical in SF and LF. The bottom row of Table 1 indicates the true average dot speed over the entire stimulus for the LF stimuli. For instance, the minimum V value is 0.011°/s in SF and 0.24°/s in LF.

3.2.1.3. Data analysis. Although the response distributions are usually not Gaussian, we chose to performed an ANCOVA on the results, so as to find the general influence of the multiple variables. We then controlled the main effect with non-parametric tests (WMP is the Wilcoxon-matched-pair test).

3.2.2. Results

3.2.2.1. Effect of the field of view (FOV) on the reported tilt sign. Fig. 8 shows the circular histograms of the bisector-oriented tilt distributions for the different W and FOV, grouping all slants, speeds and subjects, before correction for the 180° tilt ambiguity. The black arrows indicate the stimulus tilt (downward arrow) and frontal translation (upward arrow). The zero (rightward) direction indicates the bisector of these 2 directions. This graph illustrates the weight of depth reversals in the subjects' responses. As in experiment 1, depth reversals are more numerous in SF (35%) than in LF (0.1%). These percentages are both smaller than in Experiment 1. This is due to an effect of practice for subjects YF and YS who had already contributed to Experiment 1. Indeed their tilt reversals were fewer (25% and 31% in SF and 0% in LF) than the numbers found for the 2 new subjects (38 and 44% in SF, 0.4 and 0.04% in LF). We verified that the results described below did not differ between the trained subjects YF and YS, and the 2 naïve subjects.

3.2.2.2. Effect of the variables on the absolute tilt error. Fig. 4B plots the average absolute tilt error as a function of W in SF and LF. The curves are highly similar to Fig. 4A, showing a good quantitative agreement between our 2 experiments. First, the tilt errors are smaller in LF than in SF (average difference 12°, ranging between 0° and 33°, MWU test $Z = 32$, $p < 0.001$), and this is significant for each subject (MWU test at $p < 0.01$). Second, the absolute tilt error increases significantly with W for all subjects in SF, and also, to a weaker extent, in LF. For each subject, the corresponding Spearman correlation ranges between 0.42 and 0.54 ($p <$ 1E-6) in SF, and between 0.07 and 0.20 ($p < 0.001$) in LF.

In order to quantify the role of slant and dot speed on the absolute tilt error, we perform an ANCOVA with the field-size (2 levels), and W (5 levels) as factors, and slant S and speed level V as continuous predictors.

The results (Table 2A) indicate that field size and W have the most significant effects, with the significant

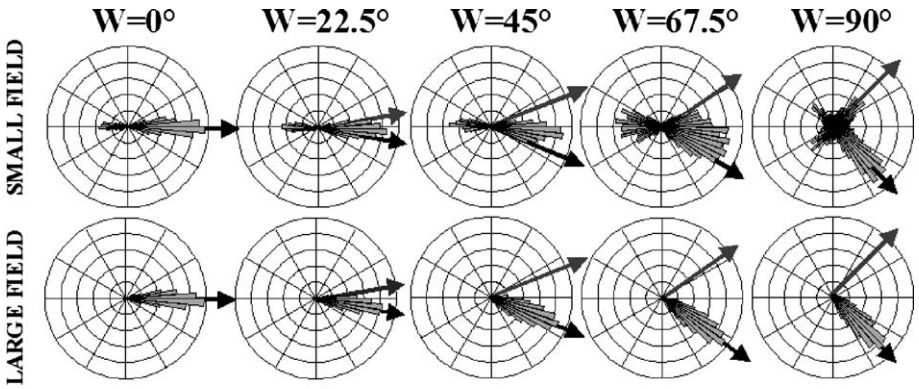

Fig. 8. Experiment 2. Circular histograms of the grouped raw responses (uncorrected for the 180° ambiguity), for each W value in small field (upper graphs) and large field (lower graphs). The black arrow indicates the true tilt direction, the grey arrow corresponds to the direction of frontoparallel translation, and the rightward horizontal indicates the bisector of these 2 directions.

Table 2
Effects of different variables on the absolute tilt error

A. ANCOVA on absolute tilt error		B. Spearman correlation between factor and absolute tilt error	
Factor	F	Small field	Large field
Field size (F)	**83.43**		
Winding angle (W)	**21.88**	**0.93**	**0.40**
Slant (S)	2.98 ($p = 0.09$)	0.03 ($p = 0.66$)	−0.002 ($p = 0.98$)
Average dot speed (V)	**48.48**	**−0.14**	**−0.15**
$F*W$	**12.51**		
$W*V$	**3.68**		

(A) Results of an ANCOVA using FOV and W as factors, and the average dot speed and the slant as covariates. (B) Spearman rank correlation coefficients between the absolute tilt error and the variables W, S, and V in small-field and large-field.

interaction just described. The dot speed V is a more influential factor than the slant. As V increases, the tilt error decreases significantly for each FOV (Table 2B, Fig. 9A). This is confirmed at the individual level for 3 of the 4 subjects in small field, and for all subjects in large field (grouped responses for all W).

The Spearman correlations in Table 2B also indicate that the effect of slant is small in LF, and is not significant in SF.

3.2.2.3. Effect of the variables on the response distributions. Fig. 8 indicates that we obtain the same trends as in experiment 1, for the grouped response distributions for all subjects, dot speed and slant values. In particular

(1) In small field the mean perceived tilt is closed to **B** (slightly toward $\boldsymbol{\tau}_s$), and the overall response variability increases with W.
(2) In large field the mean perceived tilt is closed to $\boldsymbol{\tau}_s$ (we observed a deviation of less than 1° from $\boldsymbol{\tau}_s$ for each subject and W), and the response variability increases slightly with W.

Since slant does not affect much the results, we focus on the influence of V. In small field, an increase in V "pushes" the mean of the grouped responses toward the true stimulus tilt (Fig. 9B left). In large field, the mean of the response distributions varies little with V (Fig. 9B right); for each W–FOV–slant or W–FOV–V category, this mean is less than 3.9° from the true tilt. In all conditions, we observed that the width of the distribution peaks was narrower as speed increases.

Hence, the negative correlation between the absolute tilt error and dot speed is due to a general decrease of the response variability as speed increases, coupled with a gain of accuracy in small field.

3.2.2.4. Two-peak Laplace fittings of the response distributions. We tested whether the response distributions could be the sum of two Laplace distributions, when $W = 90°$ in small field. Fig. 10 shows that the response distribution presents 2 rather symmetric peaks at less than 3° from $\mathbf{T_f}$ and $\boldsymbol{\tau}$ for 3 subjects (TS, LO, and YF). For subject YS, the distribution is flatter, with 2 underlying peaks close to the stimulus tilt. Hence, the general 2 peak shape observed at $W = 90°$ in small field (Fig. 8, upper right graph with the uncorrected 180° tilt ambiguity) is confirmed here for 3 of the 4 subjects.

Since V has a significant effect on the tilt error, we examined its effect on the grouped response distributions for small field and $W = 90°$. We found that the 2 peaks exist at each speed level, with similar locations (the tilt and the frontal translation). An increasing speed leads to the narrowing of both peaks, and the elevation of the height of the peak located at $-W/2$ (true tilt) relative to the size of the peak located at $W/2$ (frontal translation).

3.2.2.5. Separation of the effects of dot speed and 3D rotation amplitude. The previous analysis exaggerates the role of dot speed as a factor in the tilt responses for 2 reasons. First, the slant tangent varies by a factor 2 for a given V value, while V undergoes a sixfold variation for each

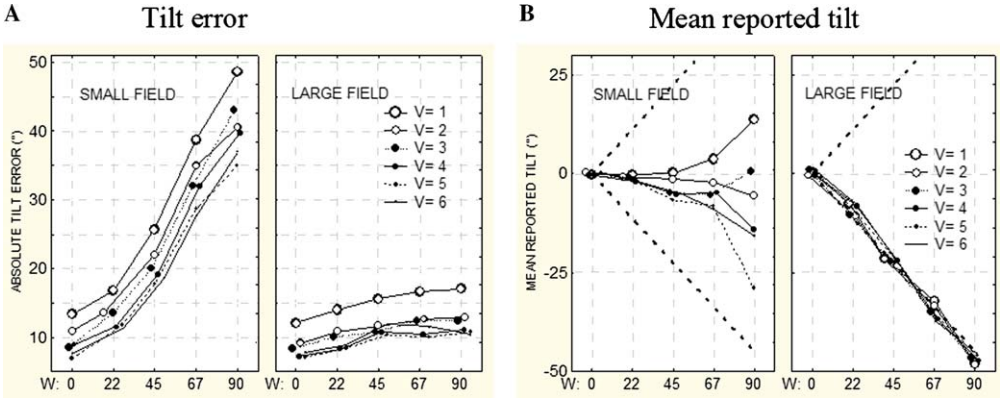

Fig. 9. Experiment 2. Analysis of the grouped tilt reports, (corrected for the 180° ambiguity) as a function of the winding angle, for different values of the dot speed, labelled from 1 to 6 (see text) in small field (left graphs) and large field (right graphs). (A) Absolute tilt error. (B) Mean reported tilt.

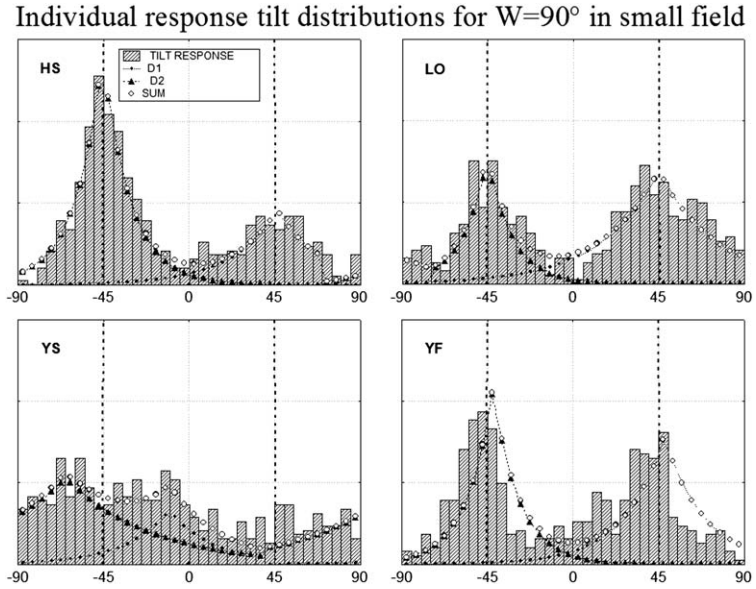

Fig. 10. Experiment 2. Individual tilt response distributions (corrected for the 180° ambiguity) for $W = 90°$ in small field, with the two-peak Laplacian fitting. D1 and D2 are the 2 fitted Laplacian functions, with their sum SUM.

slant value (Table 1). Second, for a constant slant value, V and Ra (the 3D rotation amplitude) covary. As indicated in the computational analysis below (see also Annex B), the second-order optic flow increases with Ra, and should play a key role in disambiguating the tilt direction.

In order to separate the influence of Ra and V, we measured the absolute tilt error ATEcirc from the tilt response distributions through circular statistics. This was done for each V, slant, FOV, and W value. Each distribution contained 40 responses (4 subjects, 10 tilt values). The absolute difference between the mean (by circular statistics) of this distribution, and the stimulus tilt was equal to ATEcirc. We first checked that an ANCOVA, designed as the ANCOVA above, led to similar results, in terms of the influence of V vs slant. Then we performed 2 different analyses.

First, in order to quantify the influence of Ra when V is constant, we compare ATEcirc for the minimal slant of 17.5° (i.e., for the maximum Ra value) and the maximal slant of 35° (i.e., for the minimum Ra value), in each of the 6 V-categories. Over the 30 points obtained (6 V-values, 5 W-values), the improvement of performance was significant in SF (ATEcirc decreases by 1.71°, WMP $Z = 2.44$ $p = 0.015$) but not in LF (decrease of ATEcirc 0.09°, WMP $Z = 0.91$ $p = 0.36$). Hence, Ra modulates the tilt percept in SF, but not in LF.

In a second analysis, we chose the pairs of stimuli in Table 1 that correspond to the same Ra value (up to ±10%). For instance we paired the stimulus 1 with $V_1 = .87°/s$, $S_1 = 25°$ (as dot speed and slant, respectively) with the stimulus 2 with $V_2 = 1.28°/s$ and $S_2 = 35°$, because they both correspond to a 3D rotation amplitude of 6°. Over 32 such independent pairs, we calculated the variations of ATEcirc when

(1) Ra is constant and V increases (from stimulus V_1S_1 to stimulus V_2S_2).
(2) V and Ra increase at the same time (from stimulus V_1S_1 to stimulus V_2S_1).

We also imposed that the mean relative variations of Ra and V were of similar magnitude (about 35%) for all our comparisons, between pairs of indices 1 and 2.

In SF, ATEcirc hardly decreases in the first case (mean 0.23°) and this decrease is not significant ($Z = 0.77$, $p = 0.44$). On the opposite, in the second case, the error decreases by 0.90° in average, which is significant ($Z = 2.3$, $p = 0.44$). In LF, comparisons (1) and (2) led to non-significant effects.

These two results confirm that the major effect of dot speed found above is partly due to the concomitant increase in Ra, in SF, and that maintaining a constant Ra weakens the effect of a V increase.

3.3. Conclusion

Our psychophysical results can be summarized as follows:

(1) Tilt reversals (corresponding to a 180° ambiguity on the tilt direction) are observed at a rate of about 35–40% in SF, and less than 3% in LF. Conclusions below are for responses corrected for this ambiguity.
(2) Tilt errors are larger in SF, than in LF, due to a loss of accuracy and precision in the responses. The mean perceived tilt lies between the stimulus tilt and frontal translation in small field, and near the stimulus tilt in large field.

(3) The absolute tilt error increases with W. This effect is strong in small field, and due to a decrease in the accuracy and precision of the responses. In large field, this effect is weak and due only to a slight decrease of the precision as W increases.
(4) In SF, for $W = 90°$, the mean reported tilt peaks at the stimulus tilt and at the frontal translation for grouped responses. This trend is not always verified at the individual level. Such ambiguity is not observed for lower W values, or in LF.
(5) The effect of slant is not significant, except in LF, where the performance improve slightly as slant increases. By contrast, tilt estimates improve significantly as dot speed increases; the response distributions show a higher accuracy and precision in SF, and a gain of precision in LF. This improvement is partly due to the increase of 3D rotation amplitude, which covaries with dot speed in our experiment.

4. Computational interpretation

Here, we examine the optic flow equations for a plane moving in the 3D space, under different hypotheses, and we characterize the corresponding solutions for 3D motion and structure. We then compare these solutions with the subjects' psychophysical responses described above.

4.1. The optic flow equations and their solutions in the planar case

Retinal images get formed through a perspective projection centred on the eye. Simplifications of the 3D structure from motion problem have been proposed in the literature by using either the orthographic projection, or the affine optic flow, neglecting the second-order derivatives of the optic flow. Therefore, we consider here 3 possible frameworks, (1) full flow in perspective projection, (2) affine flow in perspective projection, and (3) orthographic projection.

4.1.1. Perspective projection and full flow

We use the notations defined in Section 2. It is convenient here to introduce the proximity function $p = \frac{1}{Z}$ as

$$p = p_x \cdot x + p_y \cdot y + p_0, \tag{3}$$

where

$$\begin{aligned} &p_0 = 1/Z_0,\ x = X/Z,\ y = Y/Z,\\ &p_x = -Z_X \cdot p_0,\ p_y = -Z_Y \cdot p_0, \end{aligned} \tag{4}$$

τ and σ are related to p_x and p_y by

$$\begin{aligned} p_x &= -p_0 \cdot s \cdot \cos\tau.\\ p_y &= -p_0 \cdot s \cdot \sin\tau \end{aligned} \tag{5}$$

The coordinates of vector $\mathbf{N}$ can be written as $-1/p_0.(p_x, p_y, p_0)$

If the component of frontoparallel translation $\mathbf{T_f}$ is not zero, it is possible to detect that the optic flow field is due to a planar surface. For instance Droulez and Cornilleau-Pérès (1990) have shown that, in this case only, the spin variation (one of the second-order derivatives of the optic flow) is zero in all directions. Then, the optic flow field is related to the plane orientation and 3D motion through the equations (Longuet-Higgins, 1984):

$$\begin{aligned} u &= a_1 + a_2 \cdot x + a_3 \cdot y + A \cdot x^2 + B \cdot x \cdot y,\\ v &= a_4 + a_5 \cdot x + a_6 \cdot y + A \cdot x \cdot y + B \cdot y^2 \end{aligned} \tag{6}$$

where

$$\begin{aligned} a_1 &= T_X \cdot p_0 + \Omega_Y,\\ a_2 &= T_X \cdot p_x - T_Z \cdot p_0,\\ a_3 &= T_X \cdot p_y - \Omega_Z,\\ a_4 &= T_Y \cdot p_0 - \Omega_X,\\ a_5 &= T_Y \cdot p_x + \Omega_Z,\\ a_6 &= T_Y \cdot p_y - T_Z \cdot p_0,\\ A &= \Omega_Y - T_Z \cdot p_x,\\ B &= -\Omega_X - T_Z \cdot p_y \end{aligned} \tag{7}$$

If the coefficients a_1 to a_6, A, and B are estimated from the optic flow, 3D motion and structure can be recovered from (7). However, solving these non-linear equations leads to a twofold ambiguity, with an interchangeable role of vectors $\mathbf{n}$ and $\mathbf{t}$ (Longuet-Higgins, 1984), which are the unitary vectors of the normal $\mathbf{N}$, and translation $\mathbf{T}$, respectively. Therefore the computed orientation $\mathbf{n}'$ corresponds either to the true orientation $\mathbf{n}$, or to the 'spurious' solution $\mathbf{t}$.

4.1.2. Perspective projection and affine flow

In some cases, the optic flow can be approximated by its affine part, i.e.,

$$\begin{aligned} u &\approx a_1 + a_2 \cdot x + a_3 \cdot y,\\ v &\approx a_4 + a_5 \cdot x + a_6 \cdot y \end{aligned} \tag{8}$$

This affine approximation would not hold for a curved surface, which can induce large second-order flows even in small field (Cornilleau-Pérès & Droulez, 1989). The affine flow is characterized by the six coefficients (a_1 to a_6) of the optic flow equations. Recovering the 3D motion and structure from a_1 to a_6 yields an infinite number of solutions, because there are 6 equations, and 8 unknowns, which are p_x and p_y (or, alternatively, the tilt and slant), and the 6 coordinates of $\mathbf{T}$ and $\mathbf{\Omega}$ (p_0 is the unrecoverable scale factor).

First, let us note that the directions of the tilt and of the frontal translation can shifted by 180° simultaneously, without incidence on the affine problem. Indeed, the vectors $(-p_x, -p_y)$ and (T_X, T_Y) can be changed to their opposites (p_x, p_y) and $(-T_X, -T_Y)$ without affecting system (7). Hence, the tilt can only be recovered up to a 180° reflection with the affine flow.

Denoting $t_Z = T_Z/p_0$, and from the equations in a_2,a_3,a_5,a_6, in (7), we obtain

$$t_Z + (a_2 + a_6)/2 = (T_X \cdot p_x + T_Y \cdot p_y)/2, \quad (9)$$

$$\Omega_Z - (a_5 - a_3)/2 = (T_X \cdot p_y - T_Y \cdot p_x)/2, \quad (10)$$

which leads to

$$(t_Z - C_T)^2 + (\Omega_Z - C_\Omega)^2 = R^2, \quad (11)$$

where

$$C_T = -(a_2 + a_6)/2, \; C_\Omega = (a_5 - a_3)/2, \quad (12)$$

$$R^2 = [(a_2 - a_6)^2 + (a_5 + a_3)^2]/4. \quad (13)$$

Thus, the solution to the affine problem (six initial equations in system (7)) is such that the couple (t_Z, Ω_Z) belongs to the circle C of Eq. (11) and of radius R (positive) within the plane ($t_Z\Omega_Z$). Reciprocally, any couple (t_Z, Ω_Z) of the circle C defines one solution to the affine problem, yielding the tilt direction, and the direction of frontal translation (Annex A).

To summarize, the affine flow yields an infinite number of solutions for the 3D motion and tilt, the tilt being fully determined by the choice of the 3D motion, and vice versa. The slant value remains totally undetermined.

4.1.3. Orthographic projection

Many authors have used the orthographic projection as an approximation to small field perspective projection (Domini & Braunstein, 1998; Domini & Caudek, 1999; Todd & Perotti, 1999; Todd & Norman, 1991). Note that this approximation does not hold for translations in depth, even in small field, because such motion creates no optic flow in orthographic projection, whereas it yields an image expansion or contraction in perspective projection.

Under orthographic projection, the coordinates (u, v) of the image velocity are

$$\begin{aligned} u &= T_X + \Omega_Y/p_0 - (\Omega_Y \cdot p_x/p_0) \cdot x - (\Omega_Y \cdot p_y/p_0 + \Omega_Z) \cdot y, \\ v &= T_Y - \Omega_X/p_0 + (\Omega_X \cdot p_x/p_0 + \Omega_Z) \cdot x + (\Omega_X \cdot p_y/p_0) \cdot y \end{aligned} \quad (14)$$

These equations describe an affine flow which is equivalent to the affine part of the perspective flow (6), characterized by coefficients a_1 to a_6, where the unknowns (T_X, T_Y, Ω_X, Ω_Y) are replaced by (Ω_Y/p_0, $-\Omega_X/p_0$, $-T_Y$, T_X), and where $t_Z = T_Z/p_0$ is replaced by 0. Applying the reasoning used for the affine flow, a single point at $t_Z = 0$ is chosen on the circle, leading to the uniquely determined true tilt direction, up to a 180° reflection. Note that t_Z remains actually undetermined.

4.1.4. Summary

Table 3 summarizes the number of solutions for the plane orientation, for each projection. Each item of the triplets corresponds to the cases $W = 0°$, $0° < W < 90°$, $W = 90°$, respectively.

Table 3
Number of solutions for the computation of tilt from optic flow

Projection type	Number of solutions for τ ($W = 0°$,$0° < W < 90°$, $W = 90°$)
Perspective full flow	(1, 2, 2) $\mathbf{T_f}$
Perspective affine	(∞, ∞, ∞)
Orthographic	(2, 2, 2)-$\boldsymbol{\tau}$

The numbers in each triplet indicate the numbers of solutions for τ when $W = 0°$, $0° < W < 90°$, and $W = 90°$, respectively. The spurious solution among the 2 solutions is indicated in bold. For the perspective full flow approach, the 2 solutions merge as one for $W = 0°$, because $\tau = \mathbf{T_f}$ in this case.

4.2. Comparison of the different approaches with our results

First, we note that the second-order terms of the full flow are of the same magnitude as the first terms in LF, but much smaller in SF (Annex B). Fig. 11 illustrates the patterns of optic flow for $W = 0°$ and $W = 90°$, in SF and LF. Each condition (SF or LF) is associated with a single scale for the representations of all optic flows (first-order, second-order, and full flow, for $W = 0°$ and $W = 90°$). These 2 scales are such that the first-order flow is represented by equal arrows in LF and SF (4 top graphs). Note that the first-order flow pattern is similar in SF and LF. It is a pure compression for $W = 0°$, and a pure shear for $W = 90°$. Compared to the first-order flow, the magnitude of the second-order flow is small in SF but not in LF. Note that this second-order flow specifies in a unique way the direction of the 3D rotation, with an elongation of the approaching stimulus left side, and compression of the receding stimulus right side. This pattern is independent of W. As the sum of the first-order and second-order flows, the full flows differ radically in SF and LF.

In order to discuss our results, we distinguish three reference directions, namely the stimulus tilt $\boldsymbol{\tau_s}$, the direction of frontal translation $\mathbf{T_f}$, and their bisector $\mathbf{B}$. The distributions of reported tilt tend to be centered

on $\boldsymbol{\tau_s}$ in LF,

on $\boldsymbol{\tau_s}$ and $-\boldsymbol{\tau_s}$ in SF when W approaches 0°,

on a vector $\mathbf{B}_\tau$ located between $\mathbf{B}$ and $\boldsymbol{\tau_s}$ for $0° < W < 90°$ in SF, and its opposite $-\mathbf{B}_\tau$,

on $\boldsymbol{\tau_s}$ and $\mathbf{T_f}$ for $W = 90°$ in SF, and their opposite $-\boldsymbol{\tau_s}$ and $-\mathbf{T_f}$.

4.2.1. Large-field

In LF, the first and second-order optic flow components are of similar magnitude. Therefore, the full set of unknowns can be recovered from system (7), up to the ($\mathbf{n}$, $\mathbf{t}$) ambiguity, which predicts that $\mathbf{t}$ should be an alternative solution for the perceived tilt in the case $W > 0°$. However, in our experiments, the plane has a 3D translation of coordinates (T_X, T_Y, 0) (Annex B), which implies that the spurious perceived orientation $\mathbf{n}'$ should be frontoparallel, and thus associated with a perceived slant of 90°. Such a case can easily be discarded, because the plane would then appear as a single line. In addition, a slant close to 90° is

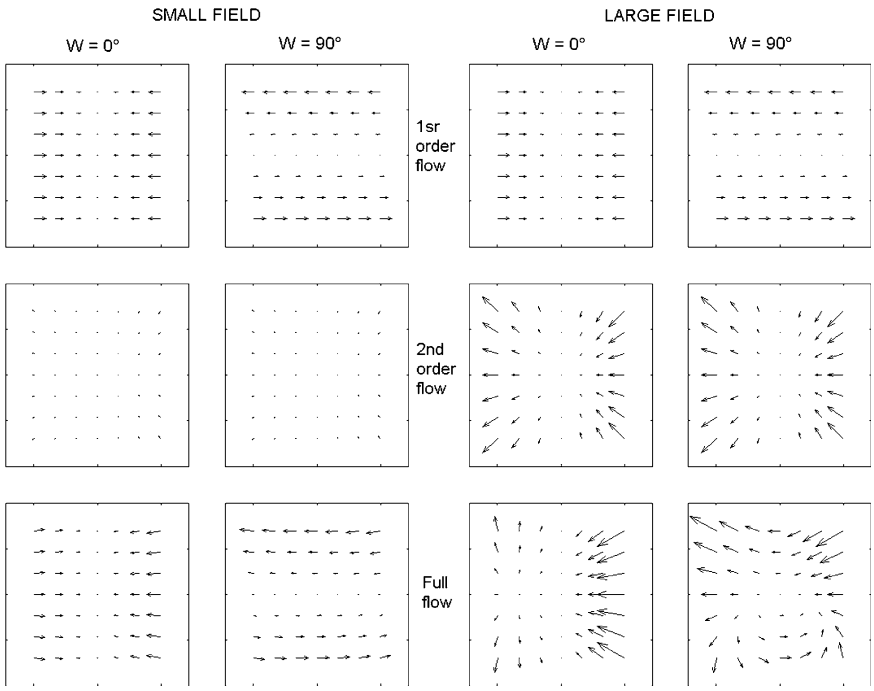

Fig. 11. Illustration of the difference in the optic flow for the SF (left graphs) and LF (right graphs) conditions, for a plane slanted in depth, rotating about a vertical axis. The tilt is horizontal for $W = 0°$, and vertical for $W = 90°$. Upper row: first-order component of the optic flow; middle row: second-order component of the optic flow; bottom row: full flow (sum of the 2 previous components). A single scale value has been used for the left graphs (SF), and for the right graphs (LF), respectively, so that the first-order flow are represented with similar arrows in the 2 conditions. The width and height of each graph are 8° for SF, and 60° for LF.

usually accompanied with strong texture gradients, blur gradients, or variations in other static depth cues, in normal vision conditions. By contrast, in our stimuli, dot density and blur are uniform within the display plane, which might lead the subjects to discard the 90° slant solution. Indeed, the results that we obtain in large field support the view that the subjects discard the case of a 90° slant, and report only the true stimulus tilt. Hence, our results in large field are well predicted by the perspective full flow model, where the solution $\sigma = 90°$ is discarded

4.2.2. Small field

In SF, the presence of a large proportion of tilt reversals support the validity of either the affine, or the orthographic scheme. The fact that the perceived tilt departs strongly from the stimulus tilt in certain configurations argues against the use of the orthographic scheme.

As far as the affine scheme is concerned, its validity is supported by the weakness of the second-order flow, relative to the first-order components (Annex B). However, the percentage of tilt reversals is only 35% in Experiment 2, which is well below the 50% predicted by a pure affine scheme. Therefore, we conclude that at 8° width, the second order terms A and B are used by our trained observers to lift partly the ambiguity on the tilt sign.

We can also observe that the magnitude of the second-order flow increases with the speed of the 3D motion (Annex B), but not with the slant. Since the speed of the 3D rotation is itself positively correlated with the image dot speed (Table 1), we deduce that the reliability of the second-order derivatives should increase with dot speed, but not with slant, which agrees with our results. Hence, this supports the view that the major difference between our LF and SF stimuli is that the former contains a more reliable second-order flow information.

Independently from the 180° ambiguity on the perceived tilt, let us examine if the affine scheme accounts for our psychophysical results in SF. First, let us note that, without any additional constraint, it predicts an infinite number of solutions, whereas the subjects' responses clearly indicate a single perceived tilt direction for $W < 90°$, and a two-fold ambiguity (although not always verified at the individual level) for $W = 90°$. In order to account for our psychophysical results, we consider here the association of the affine scheme with 3 possible constraints.

4.2.2.1. c1. The t_Z/Ω_Z constraint. A first possibility is that the visual system uses the affine scheme in association with a hypothesis on t_Z or Ω_Z within the circular constraint described above. The couple (t_Z, Ω_Z) (on the circle C) associated with the perceived tilt value would then satisfy a specific assumption. In the plane (t_Z, Ω_Z), Fig. 12 plots the circle C for the different values of W (recall that, in our experiment, we have $t_Z = \Omega_Z = 0$, which means that circle C passes through the origin). On circle C, the grey dot indicates the couples $(t_Z{}^p, \Omega_Z{}^p)$ associated with the mean reported tilt in Experiment 2, in small field. It appears that a hypothesis 't_Z minimum' accounts for the mean perceived

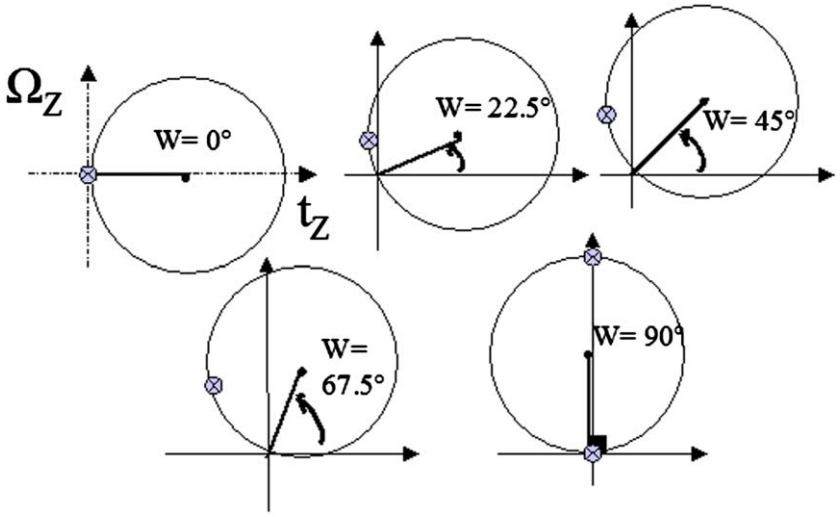

Fig. 12. Representation of the mean perceived tilt within the (t_Z Ω_Z) plane, within the affine flow approach. For each value of the angle W, for our experimental motion parameters, the grey dot indicates the position of the point of coordinates (t_Z Ω_Z) which corresponds to the mean reported tilt in the grouped responses (corrected for the 180° ambiguity) in experiment 2 in small field.

tilt when $W = 0°$ to 67.5°, because the point ($t_Z{}^p$, $\Omega_Z{}^p$) tends to satisfy this constraint. Indeed, a perceived tilt located exactly at the bisector of the tilt and frontal translation corresponds to the couple (t_Z, Ω_Z) of C, such that t_Z has the minimum (most negative) possible value on the circle. However, when $W =$ 90°, the 2 values of the reported tilt (which are close to τ and $\mathbf{T_f}$) correspond to a split between the 2 corresponding couples ($t_Z{}^p$, $\Omega_Z{}^p$). One lies at $t_Z{}^p = \Omega_Z{}^p = 0$ (the true stimulus parameters), the other at $t_Z{}^p = 0$ and $\Omega_Z{}^p > 0$.

Therefore, the affine scheme associated with the hypothesis 't_Z minimum' explains our results for $W <= 67.5°$, but not for $W = 90°$. The sudden split of the perceived tilt into 2 distinct directions at $W = 90°$ casts doubt on the validity of a constraint in t_Z or Ω_Z, associated to the affine scheme. A conclusion of this approach is also that the 3D motion associated to the perceived tilt in SF involve non-zero values of t_Z and Ω_Z for $W < 90°$, and a non-zero value of Ω_Z for $W = 90°$.

4.2.2.2. c2. The stationarity hypothesis. Wexler et al. (2001) proposed that, in presence of an ambiguity on the 3D object shape, the visual system favors the solution that corresponds to the most stationary object in the 3D space. Do our observed tilt responses in small field correspond to the minimum 3D motion of the plane dots, among the solutions of the affine problem?

For each possible solution S to the 3D affine problem (S comprises a tilt value and a 3D motion), we call $|\mathbf{U_S}|$ the norm of the 3D velocity of the object points, located on the plane P_S defined by S. The motion energy is defined as $E_S = \Sigma\ |\mathbf{U_S}|$, where the summation is carried over all image points. Annex C details the steps required to calculate for each image point the corresponding 3D coordinates of the object point on P_S, and to deduce its 3D velocity $\mathbf{U_S}$.

The stationarity hypothesis requires to find the solution S corresponding to the minimum of the function E_S. Fig. 13 presents the variations of E_S with the possible tilt direction τ_p, for the 5 values of W used in Experiment 2. The upper and lower curves are obtained for slant values of 5° and 85° (continuous lines). The curves obtained for a 25° slant are also drawn (dotted line), but are hardly discriminable from the 85° slant curves. For each stimulus configuration described by the angle W, the dashed vertical lines indicate the stimulus tilt τ (left) and the frontal

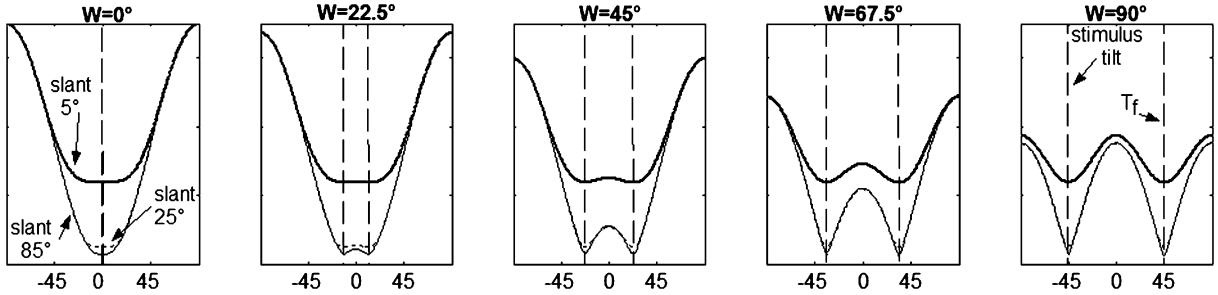

Fig. 13. Within the affine scheme, the motion energy E_S (see text) is represented as a function of the possible tilt direction, for slants 25° and 85° (two bottom superimposed curves) and 5° (upper curve). From left to right, $W = 0°$, 22.5°, 45°, 67.5°, and 90°.

translation $\mathbf{T_f}$ (right), the abscissa zero representing the bisector **B**. We observe that

(1) $E_S(\tau_p)$ tends to decrease as the slant increases from 5° to 25°, and then is relatively stationary between slants 25° and 85°
(2) whatever the slant value, the curves present 2 relative minima in τ and $\mathbf{T_f}$.

The first conclusion indicates that the variations of E_S with slant are very weak, and that a process of finding the minimum of E_S would leave slant poorly determined. The second conclusion leads to predict a $\tau/\mathbf{T_f}$ ambiguity in the general case, which differs from the $\tau/\mathbf{T_f}$ ambiguity existing in the full-flow modeling. The latter approach yields a 90° slant for the spurious solution $\mathbf{T_f}$ (see section "large field" above). On the opposite, the value of the slant remains poorly determined for the present scheme, and the slant values corresponding to the 2 minima of function E_S at τ and $\mathbf{T_f}$ are similar.

Therefore, from a purely theoretical point of view, the stationarity hypothesis predicts a 2 peak (τ and $\mathbf{T_f}$) distribution for the tilt responses. However, a computational process for finding the minimum of the function $E_S(\tau_p)$ in the presence of noise might lead to a distribution centered on a direction $\mathbf{B}_\tau$ close to **B** for $W = 0$ to 67.5°. Indeed, in this case, E_S is smaller inside the angular interval $[\tau, \mathbf{T_f}]$ than outside this interval. This means that, if the probability to choose a value of the tilt was inversely related to E_S, the directions located inside $[\tau, \mathbf{T_f}]$ would be chosen more often than those outside this interval. For $W = 90°$, the directions located inside and outside $[\tau, \mathbf{T_f}]$ would have the same probability to be selected. Therefore, the stationarity hypothesis gives some qualitative account of the fact that our observed response distributions are centered near **B** for $W < 90°$, and present 2 peaks at τ and $\mathbf{T_f}$ for $W = 90°$.

Also, we observe that the slope of the E_S curves decreases as W increases. This implies that a process of finding the minimum of the E_S curves would be more sensitive to noise for large W values, which could explain the corresponding increased variability of the tilt responses.

5. Discussion

5.1. The psychophysics of tilt perception from motion

Our results confirm the strong influence of field size and winding angle W on the ability to report the tilt of a moving plane. They are in good quantitative agreement with a previous study (Cornilleau-Pérès et al., 2002), despite a large difference in number of views displayed (72 in their study, 2 in ours), in terms of the average error in the tilt reports (Table 4).

Our results also agree with those obtained by Domini and Caudek (1999) in spite of the difference in number of views (5–83 in their studies). The tilt reports found by these authors show a mean error that can be estimated at around 10°–15° from their Fig. 12. This is smaller, for similar conditions (small-field, all W confounded) than the mean of 25° that we find in Experiment 2. This could be due to the choice of the direction of the rotation, which is random in our experiment, and fixed in theirs. In the latter case the tilt perception might be improved by an a priori knowledge (built up during experimental sessions) of the motion direction.

Using dihedral angles seen under orthographic projection, Todd and Perotti (1999) find errors of a few degrees in tilt reports, which are much lower than our observed tilt errors. Again, the use of a single motion direction in their experiment can be a factor, as well as he difference in the number of views (2 in ours and 24 in theirs). Also, these authors use a static stereoscopic view of a plane as a reference to record subjects' responses, which can itself convey hidden anisotropies of tilt perception (stereoscopy is equivalent to a motion configuration, where one view is obtained from the other by a global rotation in depth of the visual scene, involving a horizontal frontoparallel translation).

As compared to these previous studies, we do not only measure the performance in tilt reports, but we also analyze in details the shape of the response distribution. Our previous results (Cornilleau-Pérès et al., 2002) strongly suggest that our current two-view results also apply to multiple-view sequences.

In our experiments, the amplitude of the tilt variation during the motion increases with W. Could this cause the increase in the tilt error with W? As stated in the methods, the tilt direction is fixed for $W = 0°$. For $W = 90°$, it oscillates with an amplitude of ±1.5° in Experiment 1, and ±2.62° in average in Experiment 2 (maximum ±6.64°). For large-field stimuli, this increase is comparable to the growing of the absolute tilt error from 8° at $W = 0°$, to 10–15° at $W = 90°$ (Fig. 4). On the opposite, for small field stimuli, the tilt error increases by 12° to 35°–45° when W increases from 0° to 90°, which is a much larger effect than the increase of the tilt motion during a single trial. Also, the analysis of the response distributions indicates systematic

Table 4
Average absolute tilt error in this study, and in Cornilleau-Pérès et al. (2002) results

Average absolute tilt error (deg)		Small field			Large field		
		$W1$	$W2$	$W3$	$W1$	$W2$	$W3$
Our results $\sigma = 35°$		13.8	23.7	40.41	10.7	14.7	14.53
Cornilleau-Pérès et al. (2002)	$\sigma = 30°$	11	19.5	45	12	17	23
	$\sigma = 45°$	13	19	34	9	12.5	16

$W1$: $W = 0°$–$30°$, $W2$: $W = 30°$–$60°$, $W3$: $W = 60°$–$90°$.

changes in the pattern of tilt reports with W in small field, which are not simply quantitative. We conclude that the larger tilt oscillations, which occur during trials of large W values, do not account for the increase of the tilt error with W in small field.

When $W = 0°$, the first-order optic flow is a pure compression (a variation of the image velocity u along itself, represented by the term a_2 in (B3)). For $W = 90°$, it is a pure shear (a variation of the image velocity u along the orthogonal to u, represented by the term a_3 in (B3)). The visual sensitivity to compression and shear differ (e.g., Nakayama, Silverman, MacLeod, & Mulligan, 1985), and the perceived slant increases with the shear component (Domini & Caudek, 1999, fig. 18). A simple strategy for detecting the tilt direction in our stimuli would be to indicate the vector of coordinates (compression, shear). Our results demonstrate that this strategy is not applied, because when $W = 90°$ in small field, the subjects report either the true stimulus tilt parallel to the vector (compression, shear) with compression=0, but also the orthogonal direction (shear, compression). Therefore, a simple 2D scheme based on the detection of compression and shear is unlikely here.

The subjects' verbal reports indicate that the planes tended to be perceived as slightly curved (generally convex) surfaces, in both experiments in LF conditions. This is not a surprise, as our study on the estimation of surface curvature in LF demonstrates an offset of the reported curvature toward convexity for motion parallax (Cornilleau-Pérès, Tai, Cheong, & Droulez, 2005). In conditions similar to the present LF case, a concave curvature of about 0.05 to 0.09/m was perceived as zero, corresponding to a radius of 10 to 20 m for a surface tangent to a point located 1.2 m from the observer. Therefore this effect remains relatively weak. We do not know whether it interacts with tilt perception here, but the interesting point is that curvature perception is likely to involve the processing of second-order derivatives of the optic flow, which questions the interactions between curvature and orientation in 3D space perception. Pilot results showed similar tilt response distributions for a plane and for a curved surface in small-field. Thus, orientation and curvature could possibly be processed separately from the optic flow, which opens new perspectives for adapting our theoretical interpretation to the case of curved surfaces.

5.2. The roles of dot speed and rotation amplitude

In our experiments, the average dot speed is faster in large field than in small field. This is due to our choice of keeping the 3D parameters (3D rotation amplitude and slant) identical in the SF and LF conditions. Since tilt accuracy improves with dot speed (within our Experiment 2), the question arises whether this factor explains the better accuracy observed in LF, as compared to SF.

Several arguments suggest that this is not the case. First, the strong influence of W observed in SF but not in LF, indicates that dot speed (which is independent of W) is not the only key factor here. Second, our computational analysis shows that the first-order flow conveys ambiguous tilt information, and that the second-order flow disambiguates the tilt sign and direction. For a rotation in depth, the second-order derivative of the flow is proportional to the rotation amplitude Ra (up to a tangent transform, see Annex B). The influence of this factor per se (i.e., independently of dot speed) is supported by the fact that Ra, contributes to the decrease of the tilt error in experiment 2, in SF. The fact that we observe an intermediate ratio of tilt reversals in SF (35% in Experiment 2) and no such tilt reversals in LF indicates that the values of the second-order flow are close to detection threshold in SF, and well above threshold in LF. Finally, it should be noted (from Annex B) that the maximum dot speed, as proportional to the radius of the field of view, increases by about 10 times from SF to LF. The corresponding factor is 68 for the second-order flow, because it increases with the square of the field radius. Taken together, these arguments support the hypothesis that dot speed and Ra both account for the better performance in large-field.

The definitive quantitative comparison between the roles of dot speed and second-order flow, as differentiating the SF and LF tilt perception, would request the independent manipulation of the first and second order flow components, through the use of false perspective ratios (for instance stimuli calculated in parallel projection involve a zero second-order flow, and can be presented in LF). This was out of the scope of the present study, where we use the natural perspective projection in all conditions. Yet, the results of some pilot experiments confirmed that the tilt direction could be unambiguously perceived for $W = 90°$ in SF when the simulated viewing distance was artificially shortened, without increasing the average dot speed.

5.3. Relationship with the modeling of optic flow

The results in large field are well predicted by the second-order full flow modeling with an additional constraint on the slant (the slant should not approach 90°). However, as stated above, the increase of the variability of the responses with W could have a theoretical origin, or could be merely the consequence of the tilt oscillation during the plane motion.

In small field, the second-order image velocity is much smaller than the first-order flow. However, the fact that 35% (rather than 50%) errors on the tilt sign occur for the small field configuration suggests a partial use of the second-order derivatives in this case. This is in agreement with Dijkstra et al. (1995) who find that the curvature sign is reported above chance level in 8° stimuli.

Hence, a major conclusion of our study is that *the second-order flow is used in all circumstances*, but its low quantitative reliability in some cases leads to the use of the affine flow approximation by the visual system. Using the affine flow does not appear as an a priori hypothesis, but rather as an approximation for the small field situation, when

the second-order components are not reliable. This supports the view that the 3D percept depends on the robustness of the input signals, such as the second-order derivatives of the optic flow, an idea advocated by Eagle and Blake (1995). It is also favored by studies on slant perception indicating that the accuracy in slant reports improves as the second-order flow increases in small-field (Cornilleau-Pérès et al., 2000). Finally, a processing of the second-order flow in SF is indicated by experiments on the perception of surface curvature in small field (Cornilleau-Pérès & Droulez, 1989; Norman & Lappin, 1992).

The restriction of the optic flow analysis to its first order terms predicts no specific direction of the perceived tilt. It has to be associated to an additional constraint, or a priori hypothesis, to lead to a unique, or twofold percept. Among the different constraints associated to the affine flow approximation and tested above, the stationarity hypothesis seems to be the most promising account of the tilt response distributions in small-field. Note that van Boxtel et al. (2003) fail to account for the increase in tilt errors with W through the stationarity hypothesis. They find that the average 3D displacement of the object points does not depend on W, which is in agreement with our calculus (the minimum of the curve E_S at τ_S does not depend on W in Fig. 13). We add to their approach the description of the variations of E_S with the possible tilt τ of the plane. These variations depend strongly on W, which is, to some extent, compatible with the observed tilt response distributions.

However, the analysis presented gives only a preliminary support of the affine/stationarity scheme for the perception of tilt in small field. A complete demonstration of its validity would first request the use of stimuli producing to a pure affine flow (by using a smaller stimulus size, or a simulated smaller perspective angle in the stimulus generation), so as to exclude the possible use of the second-order flow, even with a coarse precision. Also, possible processes mediating the choice of the tilt value corresponding to the minimum of the function E_S have to be modeled and simulated on noisy data.

In this regard, our results indicate that the previous theoretical approaches of the computation of tilt from motion (e.g., Longuet-Higgins, 1984) have to be quantitatively simulated and studied in terms of their robustness to noise, so as to gain biological relevance. In parallel, some artificial vision studies have contributed to this question. For instance the recovery of the motion and structure parameters tends to have a maximal sensitivity to image noise when the 3D translation is parallel to N (Tsai & Huang, 1984; Daniilidis & Nagel, 1993), which supports the view that tilt estimates should lose accuracy when W increases.

It is of interest to note that different types of 3D motion are theoretically associated with different perceived tilt, for each W value in small field, within the affine scheme. For instance, at $W = 90°$, a perceived tilt equal to the true tilt corresponds to a zero sagittal rotation Ω_Z, whereas this component is not zero when the perceived tilt is parallel to the frontal translation (parag. c1 above). This suggests that the subjects might not perceive a unique type of 3D motion throughout the experimental sessions, namely a pure rotation in depth, but that other types of motion (rotations about the sagittal axis) can be perceived also. A further testing of our approach could involve a measurement of the perceived plane motion.

5.4. General conclusion

We find a strong influence of the field size and motion/orientation configuration (through the angle W between the plane normal and the frontal translation) on tilt perception. Our results support the relevance of the full flow approach in wide-field. In small field the second-order image velocity seems to be used to some extent. Yet, it is quantitatively inaccurate in that case, and the affine flow approach is relevant to explain the observed tilt ambiguities. The examination of different possible constraints associated with the affine flow analysis point to the need for further simulations of the stationarity hypothesis in small-field.

Acknowledgments

We thank TK Wong for helping in the software writing. This study was supported by Singapore Polytechnic (Toteboard funding), by the Singapore Eye Research Institute (Grant pg/98-05/00012), and by NUS (grant "Dynamic Vision and the Perception of the 3D space"). A part of our results have been presented in conferences (Zhong, Cornilleau-Pérès, Cheong, & Droulez, 2000; Zhong, Cornilleau-Pérès, Cheong, Yeow, & Droulez, 2001).

Appendix A. Relationship between the tilt direction and the 3D motion parameters in the affine scheme

We consider the case where the optic flow is approximated by its affine part as in Eq. (8). The notations of Section 2 (T_X, p_x etc...) characterize the true stimulus parameters. The same variables carrying a' typify the solution that is looked for. For instance, τ' is the tilt value to be determined among the solutions.

We define θ as the angle between the frontal translation and the axis X, within the frontoparallel plane: $T_X = t \cdot \cos\theta$, and $T_Y = t \cdot \sin\theta$,

Similarly, we have: $T'_X = t' \cdot \cos\theta'$, and $T'_Y = t' \cdot \sin\theta'$

We also write $t_Z = T_Z p_0$, and $t'_Z = T'_Z \cdot p'_0$.

In the affine scheme, we solve the first 6 equations of system (7). However, we do not consider the equations in a_1 and a_4, because they are linear in the 2 unknowns Ω'_X and Ω'_Y, which do not appear in the 4 other equations. Hence, Ω'_X and Ω'_Y can be calculated in a final stage, from the other variables, and from the observables a_1 and a_4.

We are left with the system:

$$a_2 = T'_X \cdot p'_x - t'_Z, \tag{A.1}$$

$$a_3 = T'_X \cdot p'_y - \Omega'_Z, \tag{A.2}$$

$$a_5 = T'_Y \cdot p'_x + \Omega'_Z, \tag{A.3}$$

$$a_6 = T'_Y \cdot p'_y - t'_Z, \tag{A.4}$$

which has 6 unknowns: (T'_X, T'_Y, p'_X, p'_Y, t'_Z, Ω'_Z). From 4.1.2, if the couple (t'_Z, Ω'_Z) is a solution to that system, then the point of coordinates (t'_Z, Ω'_Z) lies on the circle C of Eq. (11).

Let us show that, to each couple (t'_Z, Ω'_Z) on the circle C, there corresponds a set of solutions (T'_X, T'_Y, p'_X, p'_Y), where τ' and θ' are uniquely determined.

The system above is equivalent to:

$$t' \cdot \mathrm{tg}\sigma' \cdot \cos(\theta' + \tau') = a_6 - a_2, \tag{A.5}$$

$$t' \cdot \mathrm{tg}\sigma' \cdot \sin(\theta' + \tau') = -a_3 - a_5, \tag{A.6}$$

$$t' \cdot \mathrm{tg}\sigma' \cdot \cos(\theta' - \tau') = -a_2 - a_6 - 2 \cdot t'_Z, \tag{A.7}$$

$$t' \cdot \mathrm{tg}\sigma' \cdot \sin(\theta' - \tau') = a_3 - a_5 + 2 \cdot \Omega'_Z. \tag{A.8}$$

The first 2 equations imply that

$$t' \cdot \mathrm{tg}\sigma' = \sqrt{[(a_2 - a_6)^2 + (a_3 + a_5)^2]}. \tag{A.9}$$

Hence, $t \cdot \mathrm{tg}\sigma$ is an invariant of the optic flow (it is fully specified by the observables). It is also equal to $2 \cdot R$, where R is the radius on the circle C (see 4.1.2).

It follows that Eqs. (A5) and (A6) determine uniquely (up to 360°) the angle $\theta' + \tau'$ by

$$\cos(\theta' + \tau') = -(a_2 - a_6)/(2 \cdot R), \tag{A.10}$$

$$\sin(\theta' + \tau') = -(a_3 + a_5)/(2 \cdot R). \tag{A.11}$$

Hence, the angle $\theta' + \tau'$ is another invariant of the affine flow.

Analogously, for a given couple (t'_Z, Ω'_Z) on circle C, Eqs. (A7) and (A8) lead to specify uniquely the angle $\theta' - \tau'$. From $\theta' + \tau'$ and $\theta' - \tau'$, we can deduce θ' and τ', up to 180° each. Therefore, each couple (t'_Z, Ω'_Z) on C corresponds to 2 couples (θ', τ'), and ($\theta' + 180°$, $\tau' + 180°$), solutions to the affine problem. In turn, those couples allow to determine the 3D motion (θ' gives $T'_X$and T'_Y, and equations in a_1 and a_4 give Ω'_X and Ω'_Y) up to a scaling factor. The unknowns t' and $\mathrm{tg}\sigma'$ are undetermined, and are related by Eq. (A9).

Reciprocally, if τ' is known, then θ' is known, because the sum $\theta' + \tau'$ can be calculated from (A5) and (A6). Then, Eqs. (A7) and (A8) determine uniquely the couple (t'_Z, Ω'_Z).

In conclusion, for a given affine flow, choosing the couple (t'_Z, Ω'_Z) is on the circle C is equivalent to determining the value of the tilt (up to a 180° ambiguity), and vice versa.

Appendix B. Application of the affine perspective scheme to our experimental parameters

We use the notations of Section 2. The 3D velocity of an object point with coordinates (X Y Z) is

$$\begin{aligned} U &= T_X + \Omega_Y \cdot Z - \Omega_Z \cdot Y, \\ V &= T_Y + \Omega_Z \cdot X - \Omega_X \cdot Z, \\ W &= T_Z + \Omega_X \cdot Y - \Omega_Y \cdot X. \end{aligned} \tag{B.1}$$

In our experimental conditions, the 3D motion is a rotation around a frontoparallel axis, with no rotation around the Z-axis ($\Omega_Z = 0$).The rotation axis passes through the point (0 0 Z_0) (see methods), which is stationary in space ($U = V = W = 0$). This implies that

$$\begin{aligned} T_X &= -\Omega_Y \cdot Z_0, \\ T_Y &= \Omega_X \cdot Z_0, \\ T_Z &= 0 \end{aligned} \tag{B.2}$$

We can hypothesize, with no loss of generality, that the axis X is parallel to the translation, which implies that $T_Y = 0$, and $\Omega_X = 0$. We also have $\Omega_Y = \tan(\mathrm{Ra})$, where Ra is the 3D rotation amplitude.

From system (7), the optic flow coefficients are then:

$$\begin{aligned} a_1 &= a_4 = a_5 = a_6 = B = 0, \\ a_2 &= T_X \cdot p_x, \\ a_3 &= T_X \cdot p_y, \\ A &= -p_0 \cdot T_X \end{aligned} \tag{B.3}$$

Note that the second-order coefficient A is equal to $-p_0 \cdot T_X$, which in turn is equal to Ω_Y (from B2). This shows that, in our experiment, the second-order flow is directly proportional to tan(Ra).

In system (6), the first order flow in Y (affine part of v) is null, and the first order flow in X is

$$u_1 = T_X \cdot (p_x \cdot x + p_y \cdot y). \tag{B.4}$$

From Eqs. (2) and (4), this yields

$$u_1 = -T_X \cdot p_0 \cdot \mathrm{tg}\sigma(\cos\tau \cdot x + \sin\tau \cdot y). \tag{B.5}$$

In polar image coordinates e, ζ, we have: $x = e \cdot \cos\zeta$ and $y = e \cdot \sin\zeta$, which leads to

$$u_1 = -p_0 \cdot T_X \cdot e \cdot \mathrm{tg}\sigma \cdot \cos(\tau - \zeta). \tag{B.6}$$

Hence, the maximum absolute value of u_1 over the stimulus is obtained for $\zeta = \tau$, at the maximum image eccentricity $e = e_m$. This maximum value is:

$$u_{1\mathrm{m}} = |p_0 \cdot T_X \cdot e_m \cdot \mathrm{tg}\sigma|. \tag{B.7}$$

Let us now consider the second order flow. Its values in the u and v components are, respectively:

$$\begin{cases} u_2 = A \cdot x^2 + B \cdot x \cdot y = \Omega_Y \cdot x^2 - \Omega_X \cdot x \cdot y = -p_0 \cdot T_X \cdot e^2 \cdot \cos^2\zeta \\ v_2 = A \cdot x \cdot y + B \cdot y^2 = \Omega_Y \cdot x \cdot y + \Omega_X \cdot y^2 = -p_0 \cdot T_X \cdot e^2 \cdot \cos\zeta \cdot \sin\zeta \end{cases} \tag{B.8}$$

Therefore, the maximum absolute value of the 2nd order flow is reached by u_2 and is equal to

$$u_{2\mathrm{m}} = |p_0 \cdot T_X \cdot e_m^2|. \tag{B.9}$$

It follows that if $\mathrm{tg}\sigma \gg e_{\mathrm{m}}$, then $u_{1\mathrm{m}} \gg u_{2\mathrm{m}}$.

If α is half the viewing angle, we have $e_m = \mathrm{tg}\alpha$. Then, comparing u_{1m} and u_{2m} is equivalent to comparing $\mathrm{tg}\sigma$ and $\mathrm{tg}\alpha$. With our experimental parameters:

- $\mathrm{tg}\alpha$ is 0.07 and 0.58 in small field and large field, respectively,
- $\mathrm{tg}\sigma$ is 0.7 in experiment 1, and 0.32 to 0.7 in Experiment 2.

Hence, u_{1m} is 4.6 to 10 times higher than u_{2m}, in our small field stimuli, which leads us to regard the affine perspective scheme as valid in that case. In large-field, on the contrary, u_{1m} and u_{2m} are of comparable magnitude.

Appendix C. Computation of the motion energy E_S

Let us consider a solution S of the affine scheme, given by a tilt τ', a slant σ', a 3D rotation $(\Omega'_X, \Omega'_Y, \Omega'_Z)$ and a 3D translation (t'_X, t'_Y, t'_Z) with $t'_X = t' \cdot \cos\theta'$, and $t'_Y = t' \cdot \sin\theta'$, defined up to a scale factor. We show here that the 3D velocity $\mathbf{U_S}$ of the plane points can be written as a function of τ' and σ' only.

Given the affine flow, through coefficients a_1 to a_6 in system (7), we can calculate the constant R from (13), and the angle $u = \theta' + \tau'$ from (A10) and (A11)

$$R \cdot \cos u = (a_6 - a_2)/2,$$
$$R \cdot \sin u = -(a_3 + a_5)/2.$$

Given the tilt direction τ' and slant σ', Eqs. (A7) and (A8) lead to:

$$t'_Z = R \cdot \cos(u - 2 \cdot \tau') - (a_2 + a_6)/2,$$
$$\Omega'_Z = R \cdot \sin(u - 2 \cdot \tau') + (a_5 - a_3)/2.$$

Because $t' \cdot \mathrm{tg}\sigma' = 2 \cdot R$ (Annex A), we have $t' = R/\mathrm{tg}\sigma'$. If θ' is the angle of the frontoparallel translation, we have

$$t'_x = t' \cdot \cos\theta' = R \cdot \cos(u - \tau')/\mathrm{tg}\sigma',$$
$$t'_y = t' \cdot \sin\theta' = R \cdot \sin(u - \tau')/\mathrm{tg}\sigma'.$$

Finally, the equations in a_1 and a_4 of system (7) lead to

$$\Omega'_Y = a_1 - t'_x,$$
$$\Omega'_X = -a_4 + t'_y.$$

Assuming that the distance Z_0 is 1 (in any case, this distance is completely undetermined in the problem of computing the 3D structure and motion from the optical flow), the 3D translation (T'_X, T'_Y, T'_Z) is equal to the scaled translation (t'_x, t'_y, t'_z).

Therefore, within the affine scheme, the full 3D solution S can be expressed as a function of τ' and σ'.

From the tilt τ' and slant σ', we can also calculate the 3D coordinate (X, Y, Z) for each image point of 2D coordinate (x, y) with the equations of parag. 1

$$Z = (1 - x \cdot \mathrm{tg}\sigma' \cos\tau' - y \cdot \mathrm{tg}\sigma' \sin\tau')^{-1} \text{ and}$$
$$X = x \cdot Z \text{ and } Y = y \cdot Z.$$

The 3D motion components and the coordinates (X, Y, Z) can then be used to compute the 3D velocity U_S in each image point for the solution S:

$$\begin{aligned} U' &= T'_X + \Omega'_Y \cdot Z - \Omega'_Z \cdot Y, \\ V' &= T'_Y + \Omega'_Z \cdot X - \Omega'_X \cdot Z, \\ W' &= T'_Z + \Omega'_X \cdot Y - \Omega'_Y \cdot X, \end{aligned} \tag{C.1}$$

which in turn determines the motion energy

$$E_S = \sum \sqrt{(U'^2 + V'^2 + W'^2)},$$

the sum being taken over all image points coordinates x and y.

References

Beck, J., & Gibson, J. J. (1955). The relation of apparent shape to apparent slant in the perception of objects. *Journal of Experimental Psychology, 50*, 125–133.

Börjesson, E., & Lind, M. (1996). The effect of polar projection on the perception of euclidean structure from motion. *Perception & Psychophysics, 58*, 871–882.

Braunstein, M. L. (1968). Motion and texture as sources of slant information. *Journal of Experimental Psychology, 78*, 247–253.

Cornilleau-Pérès, V., & Droulez, J. (1989). Visual perception of surface curvature: psychophysics of curvature detection induced by motion parallax. *Perception & Psychophysics, 46*, 351–364.

Cornilleau-Pérès, V., & Gielen, C. C. A. M. (1996). The interactions between self-motion and depth perception in the processing of optic flow. *Trends in Neurosciences, 19*, 196–202.

Cornilleau-Pérès, V., Tai, L.C., Cheong, L.F., & Droulez, J. (2005). Visual distortions of the 3D visual space induced by motion parallax. Conference of the International Association for Mathematics and Computers in Simulations (IMACS), Paris, T4-I-88-0309.

Cornilleau-Pérès, V., Wexler, M., Droulez, J., Marin, E., Miège, C., & Bourdoncle, B. (2002). Visual perception of planar orientation: dominance of static depth cues over motion cues. *Vision Research, 42*, 1403–1412.

Cornilleau-Pérès, V., Wong, T. K., Cheong, L. F., & Droulez, J. (2000). Visual perception of slant from optic flow under orthographic projection and perspective projection. *Investigative Ophtalmology & Visual Sciences, 41*, 3820.

Daniilidis, K., & Nagel, H.-H. (1993). The coupling of rotation and translation in motion estimation of planar surfaces. IEEE Conference On Computer Vision and Pattern Recognition, New York,188–193.

Dijkerman, H. C., Milner, A. D., & Carey, D. P. (1996). The perception and prehension of objects oriented in the depth plane. I. Effects of visual form agnosia. *Experimental Brain Research, 112*, 442–451.

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

Domini, F., & Braunstein, M. L. (1998). Recovery of 3D structure from motion is neither euclidean nor affine. *Journal of Experimental Psychology, 24*, 1273–1295.

Domini, F., & Caudek, C. (1999). Perceiving surface slant from deformation of optic flow. *Journal of Experimental Psychology, 25*, 426–444.

Droulez, J., & Cornilleau-Pérès, V. (1990). Visual perception of surface curvature. The spin variation and its physiological implications. *Biological Cybernetics, 62*, 211–224.

Eagle, R. A., & Blake, A. (1995). Two-dimensional constraints on three-dimensional structure from motion tasks. *Vision Research, 35*, 2927–2941.

Freeman, R. B. (1966). Absolute threshold for visual slant: the effect of stimulus size and retinal perspective. *Journal of Experimental Psychology, 71*, 170–176.

Freeman, T. C. A., Harris, M. G., & Meese, T. S. (1996). On the relationship between deformation and perceived surface slant. *Vision Research, 36*, 317–322.

Garding, J., Porrill, J., Mayhew, J. E. W., & Frisby, J. P. (1995). Stereopsis, vertical disparity and relief transformations. *Vision Research, 5*, 703–722.

Gibson, J. J., & Carel, W. (1952). Does motion perspective independently produce the impression of a receding surface? *Journal of Experimental Psychology, 44*, 16–18.

Hoffman, D. D. (1982). Inferring local surface orientation from motion fields. *Journal of the Optical Society of America, A72*, 888–892.

Koenderink, J. J., & van Doorn, A. J. (1976). Local structure of movement parallax of the plane. *Journal of the Optical Society of America, 66*, 717–723.

Koenderink, J. J., & van Doorn, A. J. (1995). Relief: pictorial and otherwise. *Image and Vision Computing, 5*, 321–334.

Lappe, M. (2000). Computational mechanisms for optic flow analysis in primate cortex. *International Review of Neurobiology, 44*, 235–268.

Longuet-Higgins, H. C. (1984). The visual ambiguity of a moving plane. *Proceedings of the Royal Society of London B, 223*, 165–175.

Longuet-Higgins, H. C., & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London B, 208*, 385–397.

Marotta, J. J., Perrot, T. S., Nicolle, D., Servos, P., & Goodale, M. A. (1995). Adapting to monocular vision : grasping with one eye. *Experimental Brain Research, 104*, 107–114.

Nakayama, K., Silverman, G., MacLeod, D. I. A., & Mulligan, J. (1985). Sensitivity to shearing and compressive motion in random dots. *Perception, 14*, 225–238.

Norman, J. F., & Lappin, J. S. (1992). The detection of surface curvatures defined by optical motion. *Perception & Psychophysics, 51*, 386–396.

Norman, J. F., Todd, J. T., & Phillips, F. (1995). The perception of surface orientation from multiple sources of optical information. *Perception & Psychophysics, 57*, 629–636.

Stevens, K. A. (1983). Surface tilt (the direction of slant): a neglected psychophysical variable. *Perception & Psychophysics, 33*, 241–250.

Subbarao, M. (1988). *Interpretation of visual motion: a computational study. Series: research notes in artificial intelligence*. London: Pitman.

Todd, J., & Perotti, V. (1999). The visual perception of surface orientation from optical motion. *Perception & Psychophysics, 61*, 1577–1589.

Todd, J., & Norman, J. F. (1991). The visual perception of smoothly curved surfaces from minimal apparent motion sequences. *Perception & Psychophysics, 50*, 509–523.

Tsai, R. Y., & Huang, T. S. (1984). Uniqueness and estimation of three-dimensional motion parameters of rigid objects with curved surfaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 6*, 13–27.

van Boxtel, J. J., Wexler, M., & Droulez, J. (2003). Perception of plane orientation from self-generated and passively observed optic flow. *Journal of Vision, 3*, 318–332.

Waxman, A. M., & Ullman, S. (1985). Surface structure and three-dimensional motion from image flow kinematics. *International Journal of Robotics Research, 4*, 72–94.

Wexler, M., Panerai, F., Lamouret, I., & Droulez, J. (2001). Self-motion and the perception of stationary objects. *Nature, 409*, 85–88.

Zhong, H., Cornilleau-Pérès, V., Cheong, L.F., & Droulez, J. (2000). Visual encoding of tilt from optic flow : psychophysics and computational modelling. Lecture Notes in Computer Sciences. Springer Verlag. p800–816.

Zhong, H., Cornilleau-Pérès, V., Cheong, L. F., Yeow, M. G., & Droulez, J. (2001). Tilt perception from optic flow in two-view stimuli. Vision Sciences Society Annual Meeting. *Journal of Vision, 1*(3), 315 (abstract).