Detection of linear ego-acceleration from optic flow

Freya Festl   Department of Biology, University of Tübingen, Tübingen, Germany

Fabian Recktenwald   Department of Biology, University of Tübingen, Tübingen, Germany

Chunrong Yuan   Department of Biology, University of Tübingen, Tübingen, Germany

Hanspeter A. Mallot   Department of Biology, University of Tübingen, Tübingen, Germany

Human observers are able to estimate various ego-motion parameters from optic flow, including rotation, translational heading, time-to-collision (TTC), time-to-passage (TTP), etc. The perception of linear ego-acceleration or deceleration, i.e., changes of translational velocity, is less well understood. While time-to-passage experiments indicate that ego-acceleration is neglected, subjects are able to keep their (perceived) speed constant under changing conditions, indicating that some sense of ego-acceleration or velocity change must be present. In this paper, we analyze the relation of ego-acceleration estimates and geometrical parameters of the environment using simulated flights through cylindrical and conic (narrowing or widening) corridors. Theoretical analysis shows that a logarithmic ego-acceleration parameter, called the acceleration rate ρ, can be calculated from retinal acceleration measurements. This parameter is independent of the geometrical layout of the scene; if veridical ego-motion is known at some instant in time, acceleration rate allows updating of ego-motion without further depth-velocity calibration. Results indicate, however, that subjects systematically confuse ego-acceleration with corridor narrowing and ego-deceleration with corridor widening, while veridically judging ego-acceleration in straight corridors. We conclude that judgments of ego-acceleration are based on first-order retinal flow and do not make use of acceleration rate or retinal acceleration.

Introduction

Ego-acceleration

As we move through an environment, a pattern of image motion is induced on our retina that depends on the instantaneous ego-motion and the geometrical layout of the scene (Gibson, Olum, & Rosenblatt, 1955). Based on the resulting pattern of retinal flow, a number of parameters can be estimated, related either to the ego-motion of the observer, to scene geometry, or to combinations of these two.

Instantaneously, ego-motion can be decomposed into a rotational and a translational component, each of which can be described by three motion parameters or degrees of freedom (DoF). Theoretically, five of the six DoF of ego-motion can be recovered from the retinal motion field, including all three DoF of rotation and the two DoF comprising translational heading direction (Gordon, 1965; Longuet-Higgins & Prazdny, 1980; Koenderink & Doorn, 1987; Mallot, 2000). In a large body of research, both humans and animals have been shown to make use of optic flow for the estimation and control of ego-motion (W. H. Warren, Morris, & Kalish, 1988; Hildreth, 1992; M. V. Srinivasan, Zhang, Lehrer, & Collett, 1996; W. H. Warren, Kay, Zosh, Duchon, & Sahuc, 2001; for reviews, see Britten, 2008; Frost, 2010). The sixth degree of freedom of ego-motion, translation velocity, cannot be recovered quantitatively because projected translational flow in a given viewing direction depends only on the ratio of ego-motion and the distance of the feature point imaged in that direction. Therefore, an independent measurement of distance is needed to calibrate velocity estimates (Frenz & Lappe, 2005). Also, nonvisual information, such as proprioceptive or vestibular cues, can be integrated (see, for example, Harris, Jenkin, & Zikovitz, 2000; Ohshiro, Angelaki, & DeAngelis, 2011). Still, qualitative judgments as to the

presence or sign of motion in a given direction are possible and presumably contribute to the percept of vection, i.e., the sense of changing observer position (Berthoz, Pavard, & Young, 1975; W. H. Warren & Kurtz, 1992; Li, Sweet, & Stone, 2006; Palmisano, Allison, & Pekin, 2008).

A second group of perceptions from optic flow concerns information about the three-dimensional layout of the scene, i.e., motion parallax or structure from motion (Wallach & O'Connell, 1953). In this paradigm, the observer is assumed stationary and the three-dimensional structure of the environment is to be recovered. While initial theories attempted to recover object geometry from an inversion of the rules of (orthogonal or perspective) projection, more recently, theories have been proposed using local deformation of the retinal motion vector field (Domini & Caudek, 2003) or template matching (Fernandez & Farell, 2009). The distinction between ego-motion perception and structure from motion has been resolved in computer vision where algorithms have been developed solving both problems simultaneously (e.g., Luong & Faugeras, 1997).

Parameters combining ego-motion and environmental geometry have a high relevance for behavior. The best-studied of such parameters are time-to-collision (TTC, see D. Lee & Kalmus, 1980) and time-to-passage (TTP, see, for example, Kaiser & Hecht, 1995), which combine the distance to an approaching target with the observer's translational speed. Although both measurements individually would be prone to the previously mentioned problem of calibrating depth and ego-velocity, TTC and TTP themselves are not affected as the calibration factor cancels out. Other such combined measurements can be used for centering behavior in corridors (M. V. Srinivasan, 1998), moving through apertures (W. H. Warren Jr. & Whang, 1987), control of posture during walking (W. H. Warren, Kay, & Yilmaz, 1996), detecting deviations from the ground plane (Mallot, Bülthoff, Little, & Bohrer, 1991), etc. For the problem of running toward a flying ball and catching it, the so-called optic acceleration cancelation theory suggests that catchers control and effectively nullify the acceleration of the rising image of the ball (see McLeod, Reed, & Dienes, 2006; Fink, Foo, & Warren, 2009). In order to apply this strategy, information on image acceleration (or velocity change) of the ball is required.

For the question of whether or not subjects can judge linear ego-acceleration from optic flow, conflicting evidence exists. Kaiser and Hecht (1995) report that judgments of time-to-passage (TTP) are not affected by the acceleration of the observer, leading to an overestimation of TTP in positive ego-acceleration and an underestimation in ego-deceleration. Capelli, Berthoz, and Vidal (2010) report a gradual

influence of positive ego-acceleration but not of deceleration. In the judgment of accelerating object movement by stationary observers, acceleration cues are also not exploited (Benguigui, Ripoll, & Broderick, 2003).

Drivers in real or simulated cars have a sense of velocity change, which allows them to keep their perceived velocity constant. R. J. Snowden, Stimpson, and Ruddle (1998) asked drivers to keep their speed under changing visibility conditions in a driving simulator. In this study, drivers increased speed when visibility dropped. The drop in visibility was produced by adding fog to the scene, which should mostly affect image contrast. Presumably, drivers underestimate their speed at low visibility or contrast, and subsequently increase their speed in order to keep the resulting ego-motion estimate constant. While speed estimation thus is not veridical, (speed is underestimated at low visibility) it still shows that ego-motion change can be perceived. In real world driving, Owens, Wood, and Carberry (2010) found that perceived ego-motion velocity is largely unaffected by visibility conditions but is perceived rather more veridically. Estimates of ego-motion change are also used in visually controlled braking behavior (D. N. Lee, 1976; Fajen, 2005). In insects, keeping optic flow constant has been shown to be a behavioral strategy for achieving grazing landing maneuvers (M. V. Srinivasan, Zhang, Chahl, Barth, & Venkatesh, 2000). For the landing response of the housefly (leg extension), the control variable seems to be the total motion detector output consistent with forward motion, integrated over visual field position and time (Borst & Bahde, 1986).

A direct experiment on perceiving ego-acceleration from optic flow was carried out by Berger, Schulte-Pelkum, and Bülthoff (2010) who studied the integration of visual and vestibular information on a motion platform. The believability of brief forward accelerations was largest if forward pitch of the platform was combined with a visual acceleration stimulus. Again, this result indicates that some sense of acceleration must be obtained from optic flow.

Nonzero retinal acceleration occurs even in constant linear motion as an effect of perspective. Gordon (1965) suggested on theoretical arguments that this retinal acceleration does not play a role in ego-motion perception. However, direct measurements of the discriminability of acceleration and velocity change (De Bruyn & Orban, 1988; Calderone & Kaiser, 1989; R. J. Snowden & Braddick, 1991) indicate that human observers are able to judge the acceleration and deceleration of moving stimuli. Brouwer, Brenner, and Smeets (2002) tested the discrimination of accelerating and decelerating dots moving horizontally on the screen and found a discrimination threshold of about

(a) Feature-based acceleration rate approach

(b) Pixel-based acceleration rate approach.
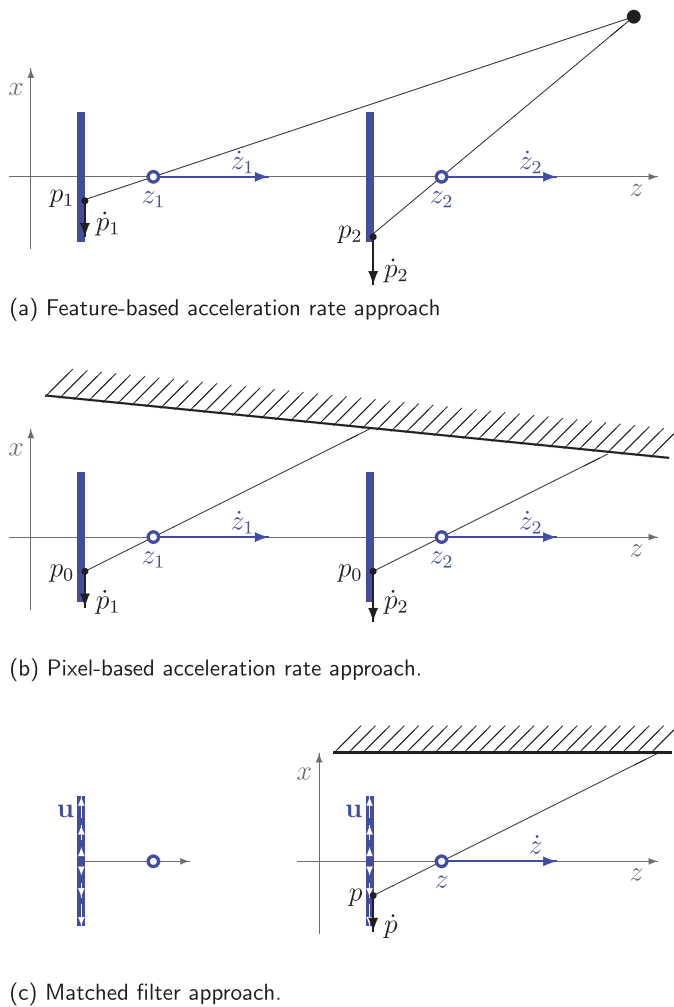

(c) Matched filter approach.

Figure 1. Three approaches to ego-acceleration estimation from optic flow. Camera frames are symbolized by an image plane (blue bars) and a nodal point (open blue circles) with position labels z, z_1, or z_2. The observer is moving along the z-axis at a speed $\dot{z}$. Image points are marked by the letter p, image motions by $\dot{p}$. (a) Two frames in the feature-based acceleration rate approach. A feature is tracked for a time long enough to measure its retinal acceleration. From this, ego-acceleration is calculated using the acceleration rate. In this approach, no assumptions about scene geometry are made. (b) Two frames from the pixel-based acceleration rate approach. Retinal motion is measured at a fixed pixel. Ego-acceleration is then estimated as before. In this approach, linear variation of lateral distance x with z must be assumed. (c) Matched-filter approach in an environment with constant depth distribution. Left: For a particular motion pattern and environment, an expected vector-field $\mathbf{u}(\mathbf{p})$ (cf. Equation 5) exists, which is symbolized in the figure by the white arrows appearing within the image plane. Right: Ego-motion and ego-acceleration can then be estimated from a comparison of the expected flow field with the actually sensed flow field.

$5°/\text{sec}^2$. However, the performance depended on the total duration of the movement. The authors therefore conclude that subjects compare initial and final motion velocity rather than actually assessing acceleration.

Neuronal responses to changing optic flow patterns generated as morphs between stationary flow fields have been studied in monkey medial superior temporal (MST) area (Paolini, Distler, Bremmer, Lappe, & Hoffmann, 2000). Only a small fraction of neurons (5 to 7%) responded specifically to field changes. The authors therefore suggest that acceleration cues are not processed in MST single unit response. For a review of optic flow-related specificities of extrastriate neurons, see Orban (2008).

Acceleration cues are also included in jitter. Palmisano et al. (2008) show that the perception of vection resulting from a forward translational optic flow stimulus is increased if the stimulus is combined with optic flow jitter simulating small observer accelerations orthogonal to the vection direction (Palmisano et al., 2008). Also, in predictive pursuit eye-movements where a stimulus temporarily vanishes behind an occluder (Becker & Fuchs, 1985), target acceleration has been shown to affect eye-position (Bennett, Xivry, Orban, Barnes, & Lefèvre, 2007). Likewise, in structure from motion, the relevance of acceleration cues has been demonstrated by Hogervorst and Eagle (2000). It therefore seems reasonable to assume that retinal acceleration pattern can be sensed and evaluated by human observers.

Theory: detecting ego-acceleration from optic flow

We restrict our analysis to pure translational movements in the forward z-direction, i.e., with the focus of expansion appearing in the center of the image. As previously discussed, the velocity of this ego-motion, $\dot{z}$, cannot be fully recovered from optic flow due to a confusion of scene-distance (depth) and velocity, which is inherent to optic flow. When traveling in a tubular corridor, fast motion in a wide corridor therefore yields the same flow field as slow motion in a narrow corridor.

This confusion also affects estimates of ego-acceleration, which is just the derivative of ego-motion. However, we will show that, for the ratio of acceleration and current velocity, $\ddot{z}/\dot{z}$, the dependence on scene geometry cancels out. In the sequel, we will refer to this quantity as "acceleration rate" and denote it with the letter ρ. Like time-to-contact, acceleration rate is a retinal parameter that can be determined from optical flow without knowing absolute depth, as it depends only on the ratio of scene distance and velocity.

Three approaches to the estimation of ego-acceleration can be distinguished (Figure 1). In the *feature-based acceleration rate approach* (Figure 1a), a feature is tracked for a period of time long enough to measure its velocity at subsequent points along its trajectory, i.e.,

different retinal locations. Acceleration rate is then computed from retinal acceleration. In the *pixel-based acceleration rate approach* (Figure 1b), optic flow at a given visual direction, or pixel, is measured at subsequent time-steps and the result is differentiated with respect to time. In this case, the subsequent motion measurements will be based on different objects or surface points as the observer is moving along. As before, acceleration rate is calculated from the retinal acceleration estimates. Finally, we consider a *matched-filter approach* (Figure 1c) in which ego-motion velocity is estimated from the flow field and ego-acceleration is obtained by simply taking the derivative of the velocity estimate without making use of retinal acceleration measurements at all.

We start by considering the feature-based acceleration rate approach. Assume an observer performing a pure translational movement in the direction of its optical axis, i.e., the z-direction. Thus, $z(t)$ is the observer's position at time t, while $\dot{z}(t)$ and $\ddot{z}(t)$ are the observer's velocity and acceleration, respectively. Consider also a feature point with coordinates $(x, y, 0)^\top$. With projection to the moving camera system, we obtain the image position of the feature point as $(p, q)^\top = 1/z(t)(x, y)^\top$. In the sequel, we will consider only the p component of image coordinates, $p = x/z$ (Figure 1a). In pure translational movement along the optical axis, $\dot{x} = 0$, and we obtain image velocity as $\dot{p} = -x\dot{z}/z^2$ and image acceleration as $\ddot{p} = -(x\ddot{z}z^2 - 2xz\dot{z}^2)/z^4$. The goal is to find an expression for observer speed change that can be calculated from image measurements alone. It turns out that this is possible for the acceleration rate defined as

$$\rho := \frac{\ddot{z}}{\dot{z}} = \frac{d}{dt}(\ln \dot{z}). \tag{1}$$

To see this, we calculate

$$\frac{\ddot{p}}{\dot{p}} = \frac{(x\ddot{z}z^2 - 2xz\dot{z}^2)z^2}{z^4 x\dot{z}} = \frac{\ddot{z}}{\dot{z}} - 2\frac{\dot{z}}{z}. \tag{2}$$

From the previous expressions for p and $\dot{p}$, it follows that $\dot{p}/p = -\dot{z}/z$. Note that $\dot{p}/p$ is the inverse of time-to-passage (Kaiser & Hecht, 1995). Substituting into Equation 2, we obtain

$$\rho := \frac{\ddot{z}}{\dot{z}} = \frac{\ddot{p}}{\dot{p}} - 2\frac{\dot{p}}{p}. \tag{3}$$

Here, p, $\dot{p}$, $\ddot{p}$ are the image position, image velocity, and image acceleration of the feature point considered. These variables can be measured from the image if the feature point is tracked over a sufficient period of time; in time-discrete systems, tracking must be achieved for at least three time frames. If $\rho > 0$, the observer is accelerating; if $\rho < 0$, the observer is decelerating. Acceleration rate ρ does not depend on x or z, i.e., it

will be the same for all possible choices of the tracked feature point. Therefore, the time course $\rho(t)$ during a longer trajectory can be measured via different feature points as long as each feature point is tracked long enough to measure its image acceleration. The quantitative value of instantaneous velocity can then be obtained by integrating Equation 1,

$$\dot{z}(t) = v_o \exp\left\{ \int_o^t \rho(t') dt' \right\}. \tag{4}$$

The parameter v_o reflects the initial calibration problem of optic flow which also remains in the acceleration rate approach. However, all subsequent ego-motion estimates can be given as multiples of this initial factor and will not require additional measurements.

In the *pixel-based acceleration rate approach* (Figure 1b), the lateral position x of the imaged object point may change as the observer moves, and the assumption $\dot{x} = 0$ used in the derivation of Equation 2 does not hold. It turns out, however, that this assumption can be relaxed to $\dot{x} \propto \dot{z}$, which is satisfied if the imaged surface is locally planar. Under this local planarity assumption, velocity estimation based on acceleration rate (Equation 4) will yield veridical results also in the pixel-based approach.

The *matched-filter approach* does not make use of retinal acceleration measurements at all, but rests on the simple idea that retinal motion vectors should be longer the faster the observer is moving. If we assume that the egocentric distribution of object distances (depth) does not change during the observer's movement, increases in retinal flow must be due to acceleration (Figure 1c). A formal scheme implementing this idea has been suggested by Franz, Neumann, Plagge, Mallot, and Zell (1999) and Franz, Chahl, and Krapp (2004). It assumes an expected flow field $\mathbf{u(p)}$ where the vectors $\mathbf{p}$ and $\mathbf{u}$ denote 2D retinal coordinates and the expected flow vector at this position, respectively. The expected flow field can be thought of as the average of all flow fields obtained from a particular ego-motion in different environments. In the case of flying in a tubular corridor, we simply use the radial pattern of unit vectors, $\mathbf{u(p)} = \mathbf{p}/||\mathbf{p}||$. An estimate of ego-velocity, v^*, can then be determined by projecting the actual retinal motions $\dot{\mathbf{p}}$ onto the locally expected flow vector $\mathbf{u} = \mathbf{p}/||\mathbf{p}||$ and summing the results over the visual field:

$$v^* = \frac{v_o}{A} \int \frac{(\mathbf{p} \cdot \dot{\mathbf{p}})}{||\mathbf{p}||} d\mathbf{p}. \tag{5}$$

Here, $(\cdot)$ denotes the dot product, $\dot{\mathbf{p}}$ is the local retinal motion, A is the area of the visual field, and the integral is taken over the entire visual field. The v^* estimator is veridical if the constant depth assumption is met. In the case of widening and narrowing corridors, it will
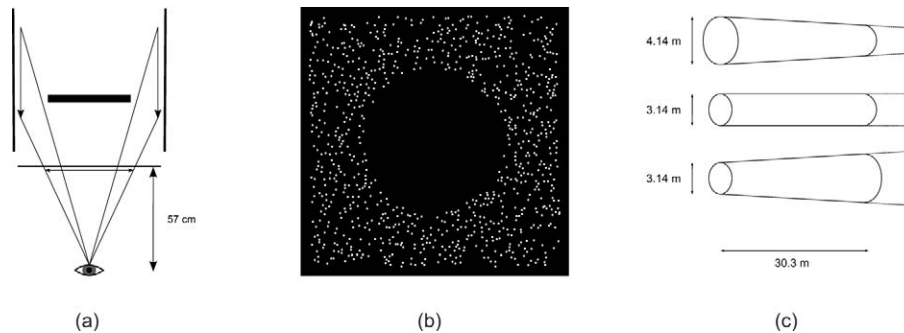
Figure 2. Apparatus and stimulus. (a) The subject was seated 57 cm in front of the screen watching the stimulus with one or both eyes. On the screen, dots corresponding to points on the corridor wall are shown. (b) Screen capture of the stimulus. For clarity, fewer and bigger dots are shown than in actual experiments. (c) The shape of the corridor was narrowing, straight, or widening.

generate a complete confusion between ego-acceleration and change of corridor diameter.

Experiment and hypotheses

In this paper, we assess subjects' ability to judge ego-acceleration in straight, narrowing, and widening corridors. In order to distinguish between pixel-based and feature-based acceleration measurements, we varied the lifetime of the dots comprising the optic flow stimulus (Sperling, Landy, Dosher, & Perkins, 1989). For short lifetimes, the feature-based mechanism should be less effective as dots can only be tracked for short times (83 ms). Dot density was high so that sufficient motion information should be available for the pixel-based mechanism at each visual direction at any time. In the long lifetime condition (1000 ms), both mechanisms can be used.

According to the three possible approaches discussed in the "theory" section and in Figure 1, three hypotheses can be formulated: (i) If the visual system uses the feature-based acceleration rate approach, veridical yes-no judgments of acceleration and corridor type should be possible. Because the feature-based approach requires tracking of image features, we predict that veridical judgments are possible only in the long dot lifetime condition, as no features can be tracked in the short lifetime condition. (ii) If the visual system uses the pixel-based acceleration rate approach, veridical yes-no judgments of ego-acceleration and corridor type should be possible both for short and long dot lifetimes. (iii) If the visual system relies on the matched-filter approach, we expect a confusion of ego-acceleration and corridor shape in both dot lifetime conditions.

The experiment was initially carried out with binocular viewing. In order to make sure that the stereoscopic cue to flatness introduced by binocular presentation did not affect the results, the measurement was repeated with monocular viewing in Experiment 2 of this study.

Methods

Participants

One male and two females (aged 25 to 30 years) participated in the experiment. One participant was one of the authors. All participants had normal or corrected-to-normal vision. Before the experiment, the participants were informed about the experimental procedure and gave their written consent on participation.

Apparatus

The visual stimuli were displayed on a HP L1950 LCD monitor with a refresh rate of 60 Hz and a resolution of 1280×1024 pixels. The stimuli had a size of 23.5×23.5 cm. The stimuli were displayed using Matlab and the Psychtoolbox. The experiment took place in a dark room. The participants were seated with their head placed in a chin rest positioned 57 cm in front of the screen, i.e., the stimuli subtended $23.3°$ of visual angle (see Figure 2a). In the binocular experiment, participants observed the stimulus with both eyes, and in the monocular experiment, the nondominant eye was covered with an eye patch.

Stimuli

The stimuli were motion sequences lasting 3 seconds (180 frames); each frame was a 800×800 pixel pattern consisting of white dots (three pixel width) on a black background. Each motion sequence simulated a flight through a cylindric or conic corridor. The dot lifetime (DLT) was either 5 or 60 frames (83 or 1000 ms). Dots were distributed randomly and homogeneously in the image, leaving out a margin of 20 pixels on all sides and a central circular region with a diameter of 480 pixels

(14°; see following). For each dot, the 3D position on the corridor wall was calculated by tracing the ray from the observer through the dot to the corridor wall. During its lifetime, the dot moved according to this 3D position. Dots reaching their lifetime limit were deleted. Deleted dots were replaced by new dots positioned randomly in the image. Due to the algorithmic procedure, there was some variance in the number of dots visible at any one time. On average, there were 1,375 ($SD = 280$) dots in each frame.

Three different corridor shapes were used (see Figure 2c). In the "straight" condition, the corridor was a cylindric tube with diameter 3.14 m. In the "narrowing" condition, the corridor was a cylindric cone with an opening (apical) angle of 2° (angle between corridor wall and axis 1°) and a diameter at starting position of 4.14 m. In the "widening" condition, the corridor was again a cone with apical angle 2° and a diameter at starting position of 3.14 m. The corridors for the widening and narrowing conditions were identical up to their orientation. The ends of the corridors were occluded by a black disk positioned at a fixed distance in front of the observer and subtending a visual angle of 14°. This occluder prevented subjects from seeing the actual end of the corridor that would have provided an additional cue to corridor shape.

For each of the three shape conditions and two dot lifetime conditions, we used 12 ego-motion profiles with constant, linear ego-accelerations of the form $z = vt + at^2/2$. Acceleration a ranged from -5.5 m/s^2 to $+5.5$ m/s^2, varying in steps of 1 m/s^2. The initial velocity v was adjusted such that the mean velocity for all stimuli was 10.1 m/s resulting in a total travel distance of 30.3 m in all cases.

Procedure

Stimuli were presented individually in random order and participants were asked to decide whether they perceived an accelerating or decelerating ego-motion (yes-no paradigm). Six psychometric curves were determined, one for each combination of the three shape conditions and the two DLT conditions. Each psychometric curve was based on 576 trials. In a combination of the method of fixed stimuli and best-PEST (Pentland, 1980), we made sure that at least 36 trials were performed with each of the 12 acceleration levels, while an additional 144 trials were assigned to the best-matching levels according to best-PEST.

Measurements were carried out in sessions of 288 trials at a time, comprising 96 trials for each of the three shape conditions. Each session lasted for about 25 minutes. The two DLT conditions were blocked with the long DLT (1000 ms) presented in a first series of six sessions and the short DLT (83 ms) in a second series of six sessions.

Data analysis

Psychometric functions were fitted to the data by a maximum likelihood procedure and the Matlab toolbox Psignifit 3.0. Psychometric functions were based on the logistic function with shift parameter α and steepness parameter β and allowed for guessing and lapsing rates γ and λ (Wichmann & Hill, 2001):

$$\Psi(x; \alpha, \beta, \gamma, \lambda) = \gamma + \frac{1 - \gamma - \lambda}{1 + \exp\{-(x - \alpha)/\beta\}}. \qquad (6)$$

The point of subjective constancy (PSC) was defined as the x-value where the participant perceived the ego-motion as constant (50% responses "accelerating," $\Psi[\text{PSC}] = 0.5$). Slopes are defined as the derivative of Ψ at PSC. Confidence intervals (CIs) of 99% were calculated at 25%, 50%, and 75% perceived acceleration, using a bootstrap sampling procedure with 2,000 samples (Fründ, Haenel, & Wichmann, 2011). The lapsing rate λ (probability that an observer misses the stimulus and decides at random) was constrained to the interval from 0 to 0.1; similarly, the guessing rate γ, was constrained to the interval from 0 and 0.01.

Theoretical predictions of the PSC shifts for the matched-filter approach were calculated according to Equation 5 on a frame-by-frame basis for all motion clips from the long DLT condition. For each corridor shape s and ego-acceleration a, this resulted in a 3-second time-course of the velocity estimator $v^*(t, a, s)$. For each shape condition, we then computed an acceleration estimate $a^*(s, a)$ as the slope of the regression line passing through the points $(t, v^*[t, a, s])$. The point of constant image flow, CIF, was defined as the simulated acceleration yielding the a^* value closest to zero:

$$\text{CIF} \ (s) = \underset{a}{\arg\min} |a^*(s, a)|. \qquad (7)$$

For the narrowing, straight, and widening corridors, CIF took the values -2.5 m/s^2, 0 m/s^2, and $+2.5$ m/s^2, respectively.

Results

Experiment 1: binocular viewing

In this experiment, the subjects viewed an optic flow pattern with both eyes. In a yes-no-paradigm, the subjects were asked if the ego-motion perception induced by the stimulus was an acceleration or a deceleration. In Figure 3, the responses are displayed as percentage of answers "acceleration." Points of subjec-
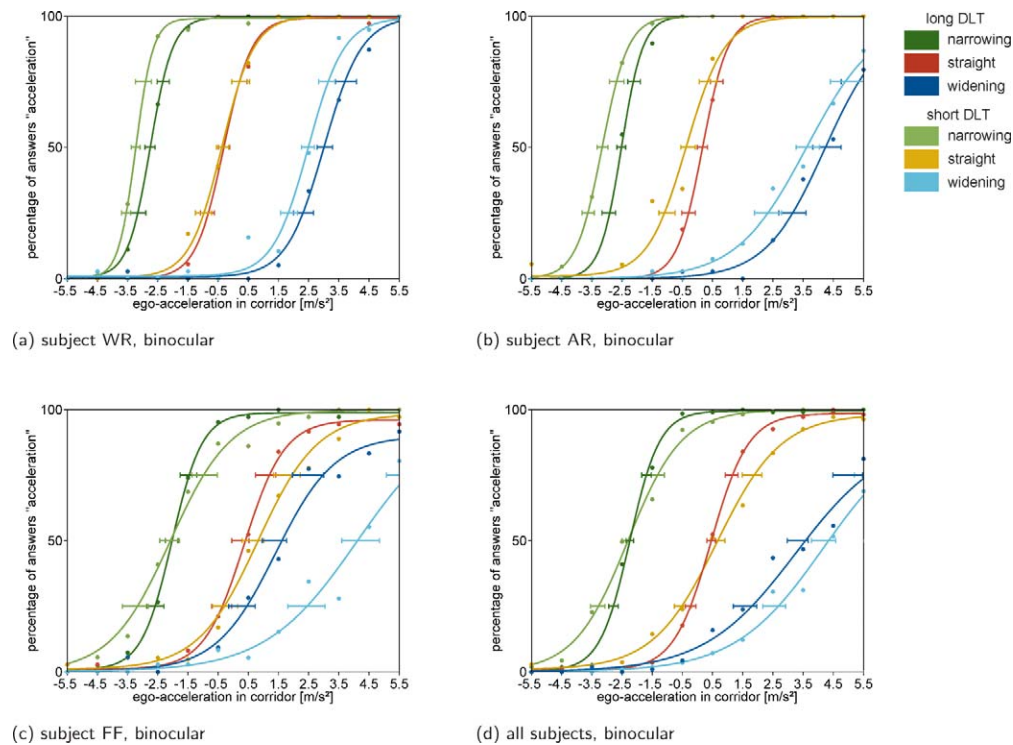
Figure 3. Psychometric functions for three subjects in Experiment 1 (binocular viewing). Green: narrowing corridor; Red, Yellow: straight corridor; Blue: widening corridor; Dark colors: long DLT; Light colors: short DLT. Horizontal "error bars" show 99% confidence intervals at the response levels 0.25, 0.5, and 0.75. The PSC values of the straight corridor are around 0 m/s^2 for all subjects. The psychometric functions of the narrowing corridor are shifted to the left, whereas the functions of the widening corridor are shifted to the right.

tive constancy of ego-motion (PSC) and slopes of the fitted psychometric functions are listed in Table 1.

In the straight corridor (yellow and red curves in Figure 3), subjects' judgments conform with the actually presented ego-acceleration values. In the narrowing corridor condition (light and dark green curves), the curves are shifted toward deceleration, indicating that, in a narrowing corridor, subjects judge decelerating stimuli as having constant speed. Vice versa, in the widening corridor (light and dark blue curves), the curves are shifted toward acceleration, indicating that, in a widening corridor, accelerating ego-motion is judged as having constant speed. The slopes of the psychometric functions do not depend on corridor shape in a systematic way.

The differences between the two DLT conditions are weaker than those between the corridor shape conditions. For the straight corridor, the confidence intervals for the two DLT conditions are overlapping, at least for subjects WR and FF. In the widening and narrowing corridors, there is only little overlap of the confidence intervals, indicating a small, but relevant effect of DLT. For subjects WR and AR, the responses for the short DLT condition are shifted consistently to the left of those from the long DLT condition. That is to say, in order to achieve the same response level, less stimulus acceleration was needed in the short DLT conditions. Note that this is not the pattern expected for the feature-based mechanism in which the displacement from zero acceleration for all curves was expected

		Narrowing corridor		Straight corridor		Widening corridor	
	Subject	Long DLT	Short DLT	Long DLT	Short DLT	Long DLT	Short DLT
PSC	WR	−2.75	−3.23	−0.32	−0.37	3.01	2.49
	AR	−2.52	−3.16	0.17	−0.37	4.21	3.62
	FF	−2.03	−2.08	0.36	0.81	1.40	4.11
Slope	WR	0.67	0.87	0.54	0.47	0.40	0.42
	AR	0.71	0.58	0.60	0.39	0.26	0.21
	FF	0.50	0.25	0.37	0.25	0.25	0.16

Table 1. Experiment 1: Points of subjective constancy (PSC) in m/s^2 and slopes of the psychometric functions in 1/(m/s^2)

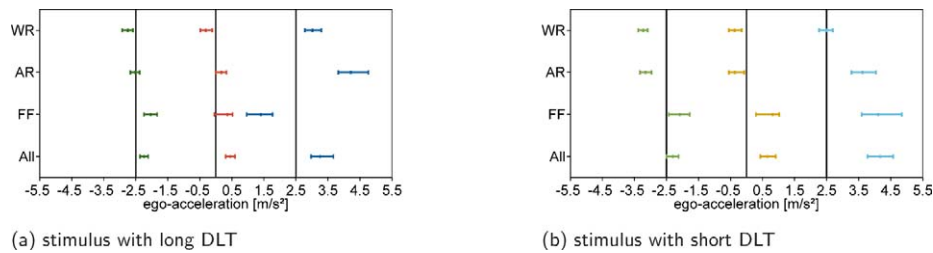(a) stimulus with long DLT   (b) stimulus with short DLT

Figure 4. PSC shifts and CIF values for Experiment 1 (binocular viewing). Horizontal lines are the 99% confidence intervals at PSC from Figure 3. Vertical lines show CIF values for the narrowing, straight, and widening corridor shape conditions (left to right). For narrowing and widening corridors, PSC values are shifted from the veridical value 0 m/s^2 toward the CIF of the according corridor.

to be larger in the small DLT condition. In subject FF, the DLT condition also affects the slope of the psychometric functions.

In order to relate the shift of the psychometric functions (PSC) to the illusory acceleration or deceleration caused by the narrowing or widening of the corridor, we used the constant image flow (CIF) values as described in the Data analysis section. Figure 4 shows these theoretical CIF values together with the shifts of the psychometric functions given by the confidence intervals at the point of subjective constancy, PSC. CIF appears to explain most of the observed shifts, both for the long and short DLT conditions.

Experiment 2: monocular viewing

In the monocular experiment, the subjects viewed the optic flow pattern with their dominant eye only. Otherwise, procedure and conditions were identical to those used in Experiment 1. As in Experiment 1, the fraction of stimuli perceived as "acceleration" was measured and a psychometric function for each of the six different conditions (three corridor shapes × two DLTs) was calculated (Figure 5, Table 2).

The results of the monocular experiment are very well in line with those of the binocular case. In particular, the pattern of PSC shift confirms the confusion of deceleration and corridor widening, and

(a) subject WR, monocular   (b) subject AR, monocular

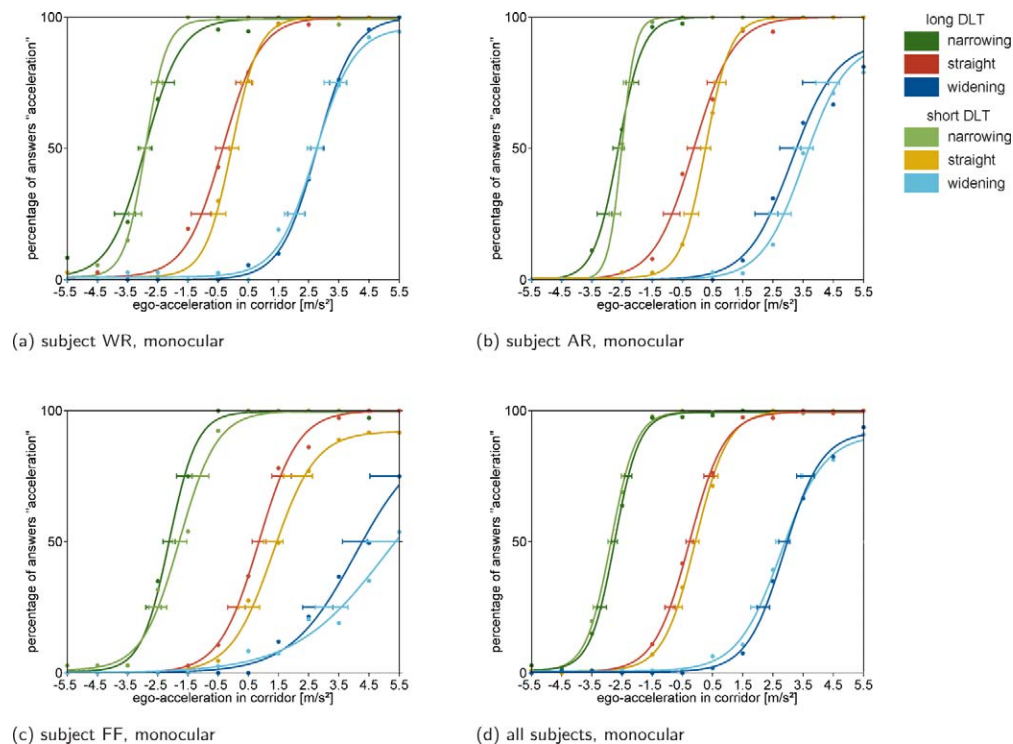(c) subject FF, monocular   (d) all subjects, monocular

Figure 5. Psychometric functions for three subjects in Experiment 2 (monocular viewing). Green: narrowing corridor; Red, Yellow: straight corridor; Blue: widening corridor; Dark colors: long DLT; Light colors: short DLT. Horizontal "error bars" show 99% confidence intervals at the response levels 0.25, 0.5, and 0.75. The PSC of the straight corridor are around 0 m/s^2 for all subjects. The psychometric functions of the narrowing corridor are shifted to the left, whereas the functions of the widening corridor are shifted to the right.

	Subject	Narrowing corridor		Straight corridor		Widening corridor	
		Long DLT	Short DLT	Long DLT	Short DLT	Long DLT	Short DLT
PSC	WR	−2.92	−2.91	−0.34	−0.04	2.78	2.73
	AR	−2.62	−2.51	−0.12	0.27	3.12	3.54
	FF	−2.13	−1.78	0.85	1.32	4.05	5.10
Slope	WR	0.42	0.72	0.40	0.56	0.41	0.39
	AR	0.62	1.13	0.38	0.61	0.34	0.36
	FF	0.54	0.40	0.38	0.36	0.23	0.16

Table 2. Experiment 2: Points of subjective constancy (PSC) in m/s^2 and slopes of the psychometric functions in $1/(m/s^2)$

the confusion of acceleration and corridor narrowing (Figure 6). In comparison to Experiment 1, the effect of dot lifetime is reduced as most of the confidence intervals are overlapping.

Discussion

The data presented in this paper clearly demonstrate that human observers do not disentangle effects of actual ego-acceleration and environment shape when judging ego-acceleration. For tubular corridors where geometry variation is absent, subjects show good discrimination of ego-acceleration. This finding is consistent with hypothesis (iii) formulated in the Introduction, stating that human observers ignore acceleration rate in spite of the fact that it would allow them to disentangle scene geometry and ego-motion change. It also shows that other schemes for simultaneous reconstruction of scene geometry and ego-motion from optic flow (Luong & Faugeras, 1997) are not used in our experiment. Variation of dot lifetime, which was introduced to favor or impede the pixel-based or feature-based algorithms, has virtually no effect.

Discrimination of ego-acceleration conditions may be based on some continuous estimate of image-acceleration or on difference calculations between image velocities at discrete times (e.g., beginning and end of the motion sequence). Some hint toward the distinction between these possibilities can be obtained

from a comparison of the performance of our subjects with measurements of just noticeable differences (JNDs) in velocity discrimination. De Bruyn and Orban (1988) measured such JNDs using planar patterns of random dots moving in unpredictable directions. Subjects had to compare two patterns displayed for 200 ms one after the other. With the stimulus parameters described in Figure 2, it can be calculated that the dot velocity in our setup is in the order of 10°/s. For this velocity, De Bruyn and Orban (1988) report Weber fractions of about 7%. This means that velocity pairs of 10°/s and 10.7°/s should be distinguishable. Thus, a just noticeable difference between the initial and the final dot velocity in our 3-second stimuli occurs for simulated ego-accelerations in the order of 0.23 m/sec^2. This is in fair qualitative agreement with our measurements where the psychometric function increases from 50% to 75% correct over a difference in ego-acceleration of about 0.5 m/sec^2. We can therefore not exclude the possibility that subjects compare subsequent velocity estimates rather than directly accessing image acceleration.

The shifts of the psychometric functions, expressed by their PSC values, are in fair quantitative agreement with the shifts predicted from the matched-filter approach to ego-acceleration estimation (Figures 4 and 6). The theoretical work on the matched-filter approach (Franz et al., 1999, 2004) was motivated by studies on the insect (dipteran) visual system where "large field neurons" have been found with receptive fields covering the entire field of view. These neurons are thought to carry out the matched-filter operation

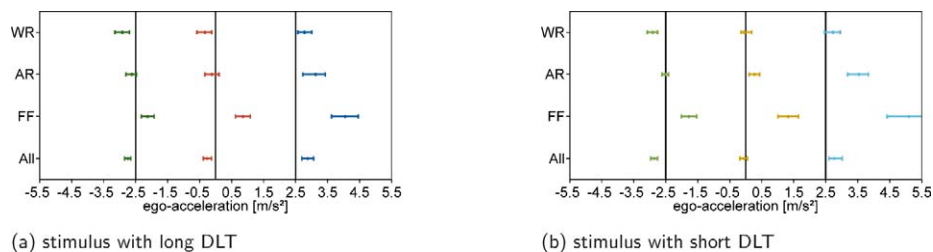(a) stimulus with long DLT

(b) stimulus with short DLT

Figure 6. PSC shifts and CIF values for Experiment 1 (binocular viewing). Horizontal lines are the 99% confidence intervals at PSC from Figure 5. Vertical lines show CIF values for the narrowing, straight, and widening corridor shape conditions (left to right). For narrowing and widening corridors, PSC values are shifted from the veridical value 0 m/s^2 toward the CIF of the according corridor.

shown in Equation 5. In primates, neurons responsible for optic flow computation have been identified in cortical area MST (e.g., Yu, Page, Gaborski, & Duffy, 2010). These neurons are thought to operate as detectors establishing a population code for various ego-motion patterns, including different heading directions. Yu et al. (2010) show that the computation performed by the receptive fields can be described as a summation over subregions with local preferred motion directions. The pattern of preferred directions corresponds to a radial motion field just as the expected motion field postulated in Equation 5. Although the function of MST optic flow neurons is generally thought to be the estimation of heading direction, they might also play a role in ego-velocity estimation. Further computational support for an optic flow scheme based on expected flow fields comes from recent work on the statistics of optic flow patterns in natural environments (Roth & Black, 2007).

The variation of DLT affects the results only marginally. In Experiment 1 (binocular viewing), psychometric functions are less sloped in the short DLT condition (Figure 3d), suggesting an increased influence of noise. In monocular viewing (Experiment 2, Figure 5d), no such effect is found. The PSC shifts found for the different subjects differ slightly between the short and long DLT conditions, but no systematic pattern is apparent.

Looking at the stimulus with one or two eyes also does not substantially affect the results. In the binocular experiment, one could argue that the dots provide stereoscopical cues to flatness that obscure the perception of the flight tunnels. If this were true, we would expect smaller PSC shifts in the monocular experiment where the 3D perception should be more veridical. The results, however, do not support this hypothesis. Rather, it seems that average optic flow velocity is evaluated with complete neglect of depth cues, be they provided by feature tracking or stereopsis.

Although the subjects in our experiments confused ego-acceleration and scene geometry, the performance is sufficient to explain subjects' ability to keep speed in driving (Owens et al., 2010) or match vestibular acceleration to visual stimuli (Berger et al., 2010). It is therefore surprising that ego-acceleration is not, or only partially used in judgments of time-to-passage (Capelli et al., 2010; Kaiser & Hecht, 1995). In the Kaiser and Hecht (1995) study, stimuli were volumes of random dots with a target dot marked by color. Capelli et al. (2010), on the other hand, used a simulated street with trees and a flag marking the target. We cannot exclude the possibility that differences of the stimulus type are responsible for the observed differences between ego-acceleration judgments and time-to-passage estimates. In natural environments, factors such as object recognition, scene segmentation, or independent motion, which have been excluded in our experiments, will also play a role.

Conclusion

In conclusion, our results show that subjects asked to judge ego-acceleration do not distinguish between effects of actual ego-acceleration and scene geometry in narrowing or widening corridors. Subjects thus confuse acceleration with corridor narrowing and deceleration with corridor widening. The mutual scaling of these two parameters is consistent with the hypothesis that ego-velocity is judged from a matched-filter process where the perceived flow field is correlated with an expected flow field for the ego-motion type in question. Acceleration cues from the retinal flow field do not enter this computation.

Acknowledgments

Commercial relationships: none.
Corresponding author: Hanspeter A. Mallot.
Email: hanspeter.mallot@uni-tuebingen.de.
Address: Department of Biology, University of Tübingen, Tübingen, Germany.

References

Becker, W., & Fuchs, A. F. (1985). Prediction in the oculomotor system: smooth pursuit during transient disappearance of a visual target. *Experimental Brain Research, 57*(3), 562–575.

Benguigui, N., Ripoll, H., & Broderick, M. T. (2003). Time-to-contact estimation of accelerated stimuli is based on first-order information. *Journal of Experimental Psychology: Human Performance and Perception, 29*(6), 1083–1101.

Bennett, S. J., Xivry, J. J. Orban de, Barnes, G. R., & Lefèvre, P. (2007). Target acceleration can be extracted and represented within the predictive

drive to ocular pursuit. *Journal of Neurophysiology, 98*(3), 1405–1414.

Berger, D. R., Schulte-Pelkum, J., & Bülthoff, H. H. (2010). Simulating believable forward accelerations on a Steward motion platform. *ACM Transactions on Applied Perception, 7*(1), 1–27.

Berthoz, A., Pavard, B., & Young, L. R. (1975). Perception of linear horizontal self-motion induced by peripheral vision (linearvection) basic characteristics and visual-vestibular interactions. *Experimental Brain Research, 23*(5), 471–489.

Borst, A., & Bahde, S. (1986). What kind of movement detector is triggering the landing response of the housefly? *Biological Cybernetics, 55*(1), 59–69.

Britten, K. H. (2008). Mechanisms of self-motion perception. *Annual Review of Neuroscience, 31*, 389–410.

Brouwer, A. M., Brenner, E., & Smeets, J. B. J. (2002). Perception of acceleration with short presentation times: can acceleration be used in interception? *Perception & Psychophysics, 64*, 1160–1168.

Calderone, J. B., & Kaiser, M. K. (1989). Visual acceleration detection: Effect of sign and motion orientation. *Perception & Psychophysics, 45*(5), 391–394.

Capelli, A., Berthoz, A., & Vidal, M. (2010). Estimating the time-to-passage of visual self-motion: Is the second order information processed? *Vision Research, 50*(9), 914–923.

De Bruyn, B., & Orban, G. A. (1988). Human velocity and direction discrimination measured with random dot patterns. *Vision Research, 28*, 1323–1335.

Domini, F., & Caudek, C. (2003). 3-D structure perceived from dynamic information: a new theory. *Trends in Cognitive Sciences, 7*(10), 444–449.

Fajen, B. R. (2005). Calibration, informatioon, and control strategies for braking to avoid a collision. *Journal of Experimental Psychology: Human Perception and Performance, 31*, 480–501.

Fernandez, J. M., & Farell, B. (2009). A new theory of structure-from-motion perception. *Journal of Vision, 9*(11):23, 1–20, http://www.journalofvision.org/content/9/11/23, doi:10.1167/9.11.23. [PubMed] [Article]

Fink, P., Foo, P. S., & Warren, W. H. (2009). Catching fly balls in virtual reality: A critical test of the outfielder problem. *Journal of Vision, 9*(13):14, 1–8, http://www.journalofvision.org/content/9/13/14, doi:10.1167/9.13.14. [PubMed] [Article]

Franz, M. O., Chahl, J. S., & Krapp, H. G. (2004). Insect-inspired estimation of egomotion. *Neural Computation, 16*, 2245–2260.

Franz, M. O., Neumann, T. R., Plagge, M., Mallot, H. A., & Zell, A. (1999). Can fly tangential neurons be used to estimate self-motion? In D. Willshaw & A. Murray (Eds.), *Icann99. Ninth International Conference on Artificial Neural Networks.* (pp. 994–999). London.

Frenz, H., & Lappe, M. (2005). Absolute travel distance from optic flow. *Vision Research, 45*(13), 1679–1692.

Frost, B. J. (2010). A taxonomy of different forms of visual motion detection and their underlying neural mechanisms. *Brain, Behavior and Evolution, 75*(3), 218–235.

Fründ, I., Haenel, N. V., & Wichmann, F. A. (2011). Inference for psychometric functions in the presence of nonstationary behavior. *Journal of Vision, 11*(6):16, 1–19, www.journalofvision.org/content/11/6/16, doi:10.1167/11.6.16. [PubMed] [Article]

Gibson, J. J., Olum, P., & Rosenblatt, F. (1955). Parallax and perspective during aircraft landings. *American Journal of Psychology, 68*(3), 372–385.

Gordon, D. A. (1965). Static and dynamic visual fields in human space perception. *Journal of the Optical Society of America, 55*, 1296–1303.

Harris, L. R., Jenkin, M., & Zikovitz, D. C. (2000). Visual and non-visual cues in the perception of linear self motion. *Experimental Brain Research, 135*, 12–21.

Hildreth, E. C. (1992). Recovering heading for visually–guided navigation. *Vision Research, 32*(6), 1177–1192.

Hogervorst, M. A., & Eagle, R. A. (2000). The role of perspective effects and accelerations in perceived three–dimensional structure-from-motion. *Journal of Experimental Psychology: Human Perception and Performance, 26*(3), 934–955.

Kaiser, M. K., & Hecht, H. (1995). Time-to-passage judgements in nonconstant optical flow fields. *Perception & Psychophysics, 57*(6), 817–825.

Koenderink, J. J., & Doorn, A. J. van. (1987). Facts on optic flow. *Biological Cybernetics, 56*(4), 247–254.

Lee, D., & Kalmus, H. (1980). The optic flow field: The foundation of vision [and discussion]. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences, 290*(1038), 169–179.

Lee, D. N. (1976). A theory of visual control of braking based on information about time-to-collision. *Perception, 5*, 437–459.

Li, L., Sweet, B. T., & Stone, L. S. (2006). Humans can perceive heading without visual path information. *Journal of Vision, 6*(9):2, 874–881, http://www.

journalofvision.org/content/6/9/2, doi:10.1167/6.9.2. [PubMed] [Article]

Longuet-Higgins, H. C., & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society (London) B, 208*(1173), 385–397.

Luong, Q. T., & Faugeras, O. (1997). Self-calibration of a moving camera from point correspondences and fundamental matrices. *International Journal of Computer Vision, 22*(3), 261–289.

Mallot, H. A. (2000). *Computational vision. Information processing in perception and visual behavior*. Cambridge, MA: The MIT Press.

Mallot, H. A., Bülthoff, H. H., Little, J. J., & Bohrer, S. (1991). Inverse perspective mapping simplifies optical flow computation and obstacle detection. *Biological Cybernetics, 64*(3), 177–185.

McLeod, P., Reed, N., & Dienes, Z. (2006). The generalized optic acceleration cancellation theory of catching. *Journal of Experimental Psychology: Human Perception and Performance, 32*, 139–148.

Ohshiro, T., Angelaki, D. E., & DeAngelis, G. C. (2011). A normalization model of multisensory integration. *Nature Neuroscience, 14*(6), 775–782.

Orban, G. A. (2008). Higher order visual processing in macaque extrastriate cortex. *Physiological Review, 88*, 59–89.

Owens, D., Wood, J., & Carberry, T. (2010). Effects of reduced contrast on the perception and control of speed when driving. *Perception, 39*(9), 1199–1215.

Palmisano, S., Allison, R. S., & Pekin, F. (2008). Accelerating self-motion displays produce more compelling vetion in depth. *Perception, 37*, 22–33.

Paolini, M., Distler, C., Bremmer, F., Lappe, M., & Hoffmann, K. P. (2000). Responses to continuously changing optic flow in area MST. *Journal of Neurophysiology, 84*, 730–743.

Pentland, A. (1980). Maximum likelihood estimation: The best PEST. *Attention, Perception, & Psychophysics, 28*(4), 377–379.

Roth, S., & Black, M. J. (2007). On the spatial statistics of optical flow. *International Journal of Computer Vision, 74*, 33–50.

Snowden, R. J., & Braddick, O. (1991). The temporal integration and resolution of velocity signals. *Vision Research, 31*(5), 907–914.

Snowden, R. J., Stimpson, N., & Ruddle, R. A. (1998). Speed perception fogs up as visiblity drops. *Nature, 392*, 450.

Sperling, G., Landy, M. S., Dosher, B. A., & Perkins, M. E. (1989). Kinetic depth effect and the identification of shape. *Journal of Experimental Psychology: Human Perception and Performance, 15*(4), 826–840.

Srinivasan, M. V. (1998). Insects as Gibsonian animals. *Ecological Psychology, 10*(3–4), 251–270.

Srinivasan, M. V., Zhang, S. W., Chahl, J. S., Barth, E., & Venkatesh, S. (2000). How honeybees make grazing landings on flat surfaces. *Biological Cybernetics, 83*(3), 171–183.

Srinivasan, M. V., Zhang, S. W., Lehrer, M., & Collett, T. S. (1996). Honeybee navigation en route to the goal: Visual flight control and odometry. *Journal of Experimental Biology, 199*, 237–244.

Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology, 45*(4), 205–217.

Warren, W. H., Kay, B. A., & Yilmaz, E. H. (1996). Visual control of posture during walking: Functional specificity. *Journal of Experimental Psychology: Human Perception and Performance, 22*, 818–838.

Warren, W. H., Kay, B. A., Zosh, W. D., Duchon, A. P., & Sahuc, S. (2001). Optic flow is used to control human walking. *Nature Neuroscience, 4*, 213–216

Warren, W. H., & Kurtz, K. J. (1992). The role of central and peripheral vision in perceiving the direction of self–motion. *Perception & Psychophysics, 51*(5), 443–454.

Warren, W. H., Morris, M. W., & Kalish, M. (1988). Perception of translational heading from optic flow. *Journal of Experimental Psychology: Human Perception and Performance, 14*(4), 646–660.

Warren, W. H., Jr., & Whang, S. (1987). Visual guidance of walking through apertures: Body-scaled information for affordances. *Journal of Experimental Psychology: Human Perception and Performance, 13*(3), 371–383.

Wichmann, F. A., & Hill, N. J. (2001). The psychometric function: I. Fitting, sampling, and goodness of fit. *Attention, Perception, & Psychophysics, 63*(8), 1293–1313.

Yu, C. P., Page, W. K., Gaborski, R., & Duffy, C. J. (2010). Receptive field dynamics underlying MST neuronal optic flow selectivity. *Journal of Neurophysiology, 103*, 2794–2807.