



Unequal retinal and extra-retinal motion signals produce different perceived slants of moving surfaces

T.C.A. Freeman *, T.A. Fowler

School of Psychology, Cardiff University, PO Box 901, Cardiff CF10 3YG, UK

Received 16 June 1999; received in revised form 29 January 2000

Abstract

Eye movements introduce retinal motion to the image and so affect motion cues to depth. For instance, the slant of a plane moving at right-angles to the observer is specified by translation and a component of relative motion such as shear. To a close approximation, the translation disappears from the image when the eye tracks the surface accurately with a pursuit eye movement. However, both translation and relative-motion components are needed to estimate slant accurately and unambiguously. During pursuit, therefore, an extra-retinal estimate of translation must be used by the observer to estimate surface slant. Extra-retinal and retinal estimates of translation speed are known to differ: a classic Aubert–Fleischl phenomenon was found for our stimuli. The decrease in perceived speed during pursuit predicts a corresponding increase in perceived slant when the eye tracks the surface. This was confirmed by comparing perceived slant in pursuit and eye-stationary conditions using slant-matching and slant-estimation techniques. Moreover, the increase in perceived slant could be quantified solely on the basis of the perceived-speed data. We found no evidence that relative-motion estimates change between the two eye-movement conditions. A final experiment showed that perceived slant decreases when a fixed retinal shear is viewed with increasing pursuit speed, as predicted by the model. The implication of the results for recovering metric depth estimates from motion-based cues is discussed. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Depth perception; Eye movements; Motion; Extra-retinal signals

1. Introduction

The perception of surface layout relies on the visual system's ability to extrapolate three-dimensional (3D) depth from two-dimensional (2D) images. Retinal images provide many useful cues to the depth, of which retinal motion is just one example. As observer and object move relative to one another, the projection of three-dimensions onto two produces informative signatures in retinal motion that correlate with the depth and surface layout in the scene. Considerable progress has been made in quantifying these signatures, revealing important methods for describing the pattern of retinal motion produced (Koenderink & van Doorn, 1976; Longuet-Higgins & Prazdny, 1980; Bruss & Horn, 1983; Koenderink, 1986; Harris, 1994). Empirical stud-

ies show the importance of retinal motion as a cue to depth (Wallach & O'Connell, 1953; Braunstein, 1968; Rogers & Graham, 1979; Braunstein & Andersen, 1981; Braunstein & Tittle, 1988; Harris, Freeman & Hughes, 1992; Meese, Harris & Freeman, 1995; Freeman, Harris & Meese, 1996; Meese & Harris, 1997). However, little attention has been paid to the role of eye movements in depth perception, which is surprising because eye movements cause wholesale changes to the pattern of retinal motion and so, in some situations, modify motion cues. How observers compensate for these changes is an issue tackled in this paper. We examine some of the consequences of eye movements on motion cues to surface slant and show that perceived slant changes in a predictable manner when observers move their eyes.

In order to understand the potential effect eye movements may have on slant perception, we first need to consider the retinal information an observer might use to judge slant when their eye is stationary. In this situation, a moving plane stimulates the retina with a

* Corresponding author. Tel.: +44-2920-874554; fax: +44-2920-874858.

E-mail address: freemant@cardiff.ac.uk (T.C.A. Freeman)

characteristic velocity gradient (Gibson, Gibson, Smith & Flock, 1959). The type of gradient depends on the tilt direction and slant of the surface together with its direction of motion. Fig. 1 depicts four examples of

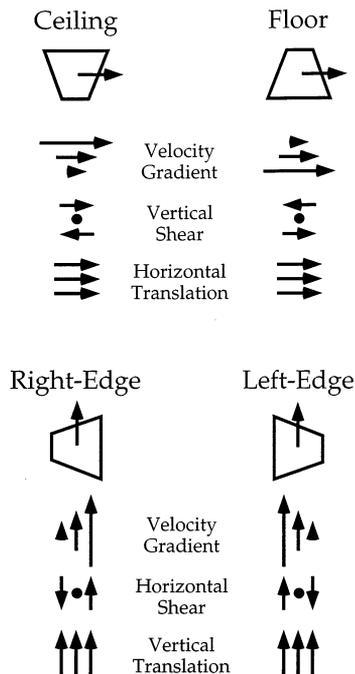


Fig. 1. Velocity gradients and component motions associated with slanted surfaces moving at right angles to the observer. In the top row, a horizontally moving 'ceiling' and 'floor' give rise to velocity gradients that are closely approximated by the sum of a vertical shear and horizontal translation. In the bottom row, a vertically moving surface with a leading 'right-edge' or 'left-edge' give rise to velocity gradients that are the sum of a horizontal shear and vertical translation.

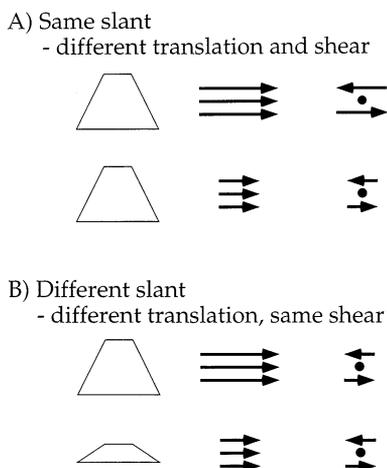


Fig. 2. Two demonstrations that translation must be used to scale the shear component. (A) Two surfaces with the same slant but moving at different speeds produce different magnitudes of shear. (B) Two surfaces with different slants moving at different speeds can produce the same magnitude of shear. In both cases, the shear must be scaled by the translation to recover slant.

moving surfaces that were used in our experiments: a 'ceiling' and a 'floor' moving horizontally, and a leading 'right-edge' and 'left-edge' moving vertically. The corresponding velocity gradients are shown below these. The gradients consist of two components of motion: translation and what we refer to as a 'relative-motion' component. For these particular surfaces, the relative-motion component is adequately approximated by a one-dimensional shear. The term 'shear' is used because the speeds that accompany this motion component are always directed at right-angles to the velocity gradient. Fig. 1 shows the shears associated with the four types of surface. By adding the shear and translation components together, the appropriate velocity gradient is produced.

To a close approximation, tracking any of these surfaces with a stabilising eye movement removes the translation component from the image.¹ Thus, during accurate pursuit, the retina is stimulated by the relative-motion component alone. This would not present too much of a problem to the observer if slant perception were based exclusively on estimating the relative-motion component. However, it turns out that both relative motion and translation are needed to judge slant accurately and unambiguously. The intuition for this is provided in Fig. 2. In Fig. 2A, two planes with the same slant are shown moving at different speeds. As speed increases, so does the magnitude of the relative-motion component; thus, estimating shear by itself would not reveal the similarity of the slants depicted. In Fig. 2B, planes of differing slant are shown. By adjusting the translation speed appropriately, the two moving planes can be made to produce the same magnitude of shear. Again, estimating slant on the basis of shear fails to discriminate the slants of the two planes.

Estimating the magnitude of slant therefore requires knowledge of both the relative motion component and translation (Koenderink, 1986; Harris, 1994; Freeman et al., 1996; Domini & Caudek, 1999). In effect, translation is used to scale the shear. This is summarised in the computational formula relating slant (σ) to shear (S) and translation (T):

$$\sigma = \arctan\left(\frac{2S}{T}\right) \quad (1)$$

One aspect of this formula that will become important is that slant is inversely related to translation speed. So for a fixed magnitude of shear, slant decreases as translation speed increases.

This observation is particularly relevant if we consider what happens when an observer makes a pursuit

¹ Eye rotations do not give rise to pure translation of the image and so this is only an approximation. For the conditions of the experiments reported here, the difference between planar translation and polar eye rotation is small and so the approximation suffices.

eye movement to track the surface. Because retinal translation is deleted by the eye movement, the visual system must rely on extra-retinal information about translation speed in order to scale the relative-motion component. One method for obtaining an extra-retinal estimate of speed is to monitor eye-velocity signals instructing the eye to move (von Holst, 1954). Intriguingly, previous work shows that moving objects appear slower when pursued. Referred to as the Aubert–Fleischl phenomenon, the difference in perceived speed is evidence that extra-retinal estimates of object speed are lower than corresponding retinal estimates (Dichgans & Brandt, 1972; Mack & Herman, 1973, 1978; Freeman & Banks, 1998; for an alternative view, see Wertheim, 1994). Demonstrating the same phenomenon for slanted surfaces predicts from Eq. (1) that perceived slant should increase during pursuit. In the experiments reported below, we first show that stimuli comprising shear and translation appear slower when pursued. We then show that their perceived slant increases during pursuit. We then consider whether changes in relative-motion estimates contribute to the increase in perceived slant. No such evidence is found, leading us to conclude that the increase in perceived slant is based entirely on the mismatch between retinal and extra-retinal estimates of object speed.

2. General methods

2.1. Stimuli

Moving random dot patterns were displayed on the black background of a Mitsubishi Diamond Pro 20 monitor at a frame rate of 100 Hz. Displays were generated by a VSG 2/3F graphics card controlled by a Pentium PC. Dot patterns had a density of ~ 0.4 dots/deg². Dot positions (x , y) were selected pseudo-randomly for the first frame of animation. Stimuli consisting of vertical shear and horizontal translation (Fig. 1, top row) were produced by updating the x component of each dot at frame-rate, using the relationship $x + sy + t$, where s and t are the components shear and translation divided by the frame rate. Stimuli consisting of horizontal shear and vertical translation (Fig. 1, bottom row) were produced by updating the y component using the relationship $y + sx + t$. Dot position was rendered with sub-pixel accuracy using an anti-aliasing technique that altered the centroid of a 2×2 pixel cluster (Georgeson, Freeman & Scott-Samuel, 1996). Each pixel subtended $\approx 0.04^\circ$ at the viewing distance of 57.3 cm. A short vertical line served as a fixation point. This was placed at the centre of an annulus window with an inner radius of 1° and outer radius of 10° . The window made peripheral and central dots invisible to the observer. All stimuli were viewed

monocularly. Any incidental visual references were made invisible by viewing the stimuli in a dark room. This prompted the use of extra-retinal estimates of object speed in the pursuit conditions of the experiments and also encouraged accurate eye movements.

Stimuli were presented with either no eye movement or an accompanying eye pursuit. In pursuit conditions, fixation point and window moved at the same velocity as the translation; in eye-stationary conditions, fixation point and window were stationary. Assuming accurate pursuit, the same retinal region was therefore stimulated in the two conditions.

During any stimulus presentation, the fixation point appeared alone for 500 ms then moved at the appropriate speed for a further 2000 ms. The random dot pattern appeared for 1000 ms following the initial 500 ms of fixation-point motion. The 500 ms of fixation-point motion that occurred before and after displaying the dot pattern was intended to promote accurate eye tracking when the dot pattern was visible. In all experiments apart from Experiment 4, the sign of translation (and therefore pursuit) was alternated from trial to trial to counteract any bias introduced by retinal or extra-retinal aftereffects. Thus, in the case of the ceiling/floor conditions, the translation alternated first leftward then rightward. The sign of the shear component was alternated in conjunction with the direction of translation to preserve the tilt direction (i.e. whether the stimulus resembled a ‘ceiling’, a ‘floor’, a ‘right-edge’ or a ‘left-edge’ surface).

2.2. Eye movement recording and analysis

In Experiments 2 and 4, eye movements were recorded using a head-mounted video-based eye tracker (Applied Sciences Laboratories Series 4000). The system was calibrated using an array of 3×3 targets at known angular deviations from the observer. Eye position was sampled at 50 Hz. Eye position records were analysed off-line in MatLab software. To obtain eye velocity, position records were first low-passed filtered and the time derivative taken. Saccades were identified in the velocity record using a velocity threshold of $\pm 70^\circ/\text{s}$. In Experiment 2, we found 1.60% trials to contain one or more saccades using this definition. The percentage was also low in Experiment 4. For this reason, pursuit accuracy was evaluated by using the eye records from remaining ‘non-saccadic’ trials, as opposed to more traditional approaches, such as replacing the saccadic region with a straight-line segment before averaging all velocity recordings (see Wyatt, 1998).

2.3. Observers

The two authors (TCAF, TAF) participated in all experiments. At least one naive observer, unaware of

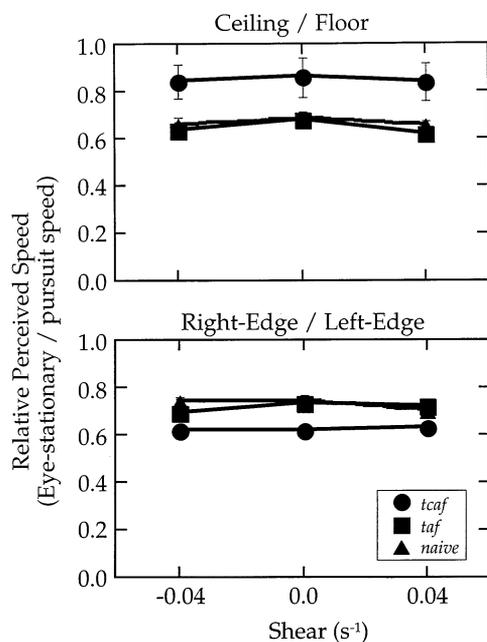


Fig. 3. Aubert–Fleischl phenomenon for shearing, translating stimuli. Data are for four observers. Different naive observers took part in the ceiling/floor and right-edge/left-edge conditions (top and bottom panels, respectively). The relative perceived-speed ratio is less than 1, indicating that moving, slanted surfaces appear slower when pursued. Error bars are ± 1 SE.

the purposes of the experiment, was used to confirm the results of the authors. In Experiments 1–3, SJMF and JHM participated in the ceiling/floor and left-edge/right-edge conditions, respectively. SJMF was an experienced psychophysical observer and JHM was not. KD, an inexperienced psychophysical observer, participated in Experiment 4 as part of her course credit requirement for the undergraduate program at the School of Psychology, Cardiff University.

3. Experiment 1: slanted surfaces appear slower when pursued

In this experiment, we sought evidence for an Aubert–Fleischl phenomenon for slanted surfaces. A speed-matching method determined the speed of eye-stationary and pursuit stimuli that yielded a perceived-speed match.

3.1. Procedure

Two stimulus intervals were shown in sequence. The pursuit interval contained a moving fixation point (and window) and the eye-stationary interval contained a stationary fixation point. The pursuit interval always appeared first. Speed matches were obtained for a fixed translation speed of $4^\circ/\text{s}$ in the pursuit interval by varying the speed in the eye-stationary interval. Follow-

ing each trial, observers indicated which interval appeared to move faster with respect to the head. The speed was then adjusted using a one-up-one-down staircase which converged on the 50% point of the psychometric function. Staircases terminated following 12 reversals; the 50% point was estimated on the basis of the mean speed over the last ten.

Ceiling/floor and right-edge/left-edge tilt directions were investigated using three magnitudes of shear: -0.04 , 0 and $+0.04 \text{ s}^{-1}$. Two staircases were assigned to each shear condition, giving a total of six staircases. These were randomly interleaved in any one session of data collection.

3.2. Results and conclusions

To examine whether these stimuli produced an Aubert–Fleischl phenomenon, the eye-stationary speeds yielding a perceived-speed match were first divided by the fixed translation speed in the pursuit interval. We refer to this ratio as the *relative perceived speed*: values > 1 imply that moving surfaces appeared faster when pursued; a value < 1 represents a classic Aubert–Fleischl phenomenon. Fig. 3 plots the results for four observers. Each data point represents the mean relative perceived speed for three replications of the experiment in the ceiling/floor and right-edge/left-edge conditions (top and bottom panels, respectively). For each observer, the relative perceived speed was consistently < 1 in each condition. Thus, moving, slanted surfaces appear slower when pursued. The mean value of relative perceived speed was 0.70, a value that agrees with previous reports for this phenomenon (e.g. Mack & Herman, 1978; Freeman & Banks, 1998). The strength of the Aubert–Fleischl phenomenon did not depend on the magnitude of shear. It was also similar for horizontal and vertical directions of pursuit.

The existence of an Aubert–Fleischl phenomenon for these stimuli predicts that their perceived slant should increase during pursuit. This hypothesis was tested in Experiments 2 and 3.

4. Experiment 2: perceived slant increases during pursuit

We investigated the slant hypothesis using a similar two-interval, forced-choice technique. This time, however, we determined the magnitude of shear that yielded a perceived-slant match between pursuit and eye-stationary stimuli.

4.1. Procedure

Slant matches were obtained using a method-of-constant stimuli. The shear in the pursuit interval was fixed

at 0.04 s^{-1} . The magnitude of shear in the eye-stationary interval was selected at random from 0.0 to 0.08 s^{-1} in 0.02 s^{-1} steps. The translation speed was $4^\circ/\text{s}$ in both intervals. Following each trial, observers selected which of the two intervals appeared ‘more slanted’. Each session consisted of 20 replications at each of the five shear magnitudes, giving 100 trials in total. An additional five practice trials were presented at the start of each session; the data from these were discarded. Psychometric functions were obtained by fitting the data with a logistic function (Maxwell, 1959) using a least-squares technique. Slant matches were defined as the 50% point on this function. Eye movements were also recorded for the two authors in the third and final replication of the experiment.

4.2. Results and conclusions

To determine whether perceived slant changed during eye pursuit, the eye-stationary shear that produced a

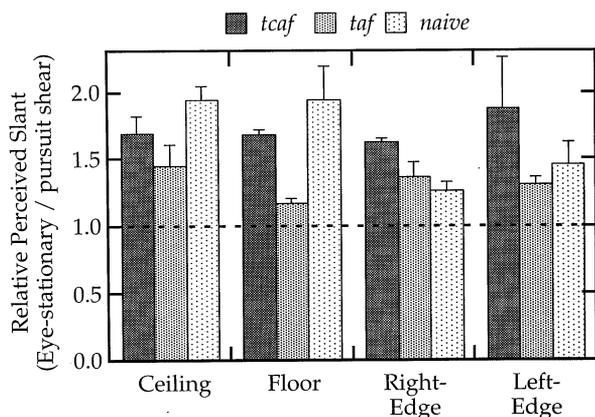


Fig. 4. Perceived slant increases when slanted surfaces are tracked by a pursuit eye movement. Data are for four observers, with different naive observers taking part in the ceiling/floor and right-edge/left-edge conditions. The relative perceived-slant ratio is > 1 , indicating that perceived slant was greater in the pursuit interval. Error bars are 1 SE.

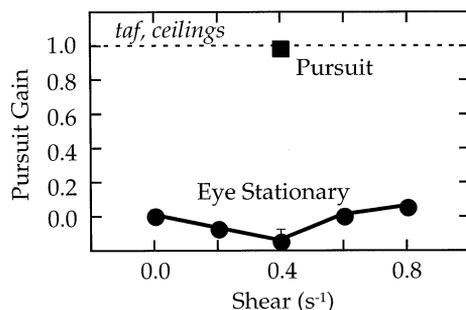


Fig. 5. Eye movement accuracy for observer TAF in the ‘ceiling’ condition of Fig. 4. Pursuit gain was close to 1 for the pursuit intervals (square symbol) and close to 0 for the eye-stationary intervals (circles). Error bars are ± 1 SE.

perceived-slant match was first divided by the fixed shear of 0.04 s^{-1} presented in the pursuit interval. We refer to this ratio as the *relative perceived slant*. A value < 1 indicates that the stimuli appeared less-slanted during pursuit. Fig. 4 plots the results for four observers. For all observers, relative perceived slant was > 1 in each condition. According to these data, therefore, perceived slant increases during pursuit. Similar findings (not shown) were obtained for a translation speed of $2^\circ/\text{s}$.

The results of the eye-movement analysis showed that eye movements were accurate in all conditions. In the analysis, the mean pursuit speeds were divided by the translation speed to obtain an estimate of pursuit gain in each condition. Fig. 5 shows the gains found for TAF in the ceiling condition. Collapsing across all tilt and pursuit directions, the mean gain in the pursuit interval was 1.09 (SE = 0.04) for this observer and 1.04 (SE = 0.04) for TCAF. In the eye-stationary intervals, there was evidence that observers made small optokinetic eye movements. On average, the mean gain was close to 0. Collapsing across both tilt direction and shear, the mean gain in the eye-stationary interval was 0.04 (SE = 0.02) for TAF and 0.06 (SE = 0.01) for TCAF.

One possible explanation of the increase in perceived slant during pursuit is that the stimuli in the eye-stationary conditions did not appear consistently slanted. We therefore repeated the experiment with stationary fixation in both intervals. Observers were perfectly able to perform the task and produced sensible psychometric functions. Another possible explanation, suggested to us by one of the reviewers, concerns the effect of moving the window in one interval whilst keeping it stationary in the other. This produces different rates of dot appearance and disappearance at the edges of the display that may have interacted with the slant percept. The edge effects may also lead to different impressions of the distance of stimuli from the observer: in the eye-stationary interval, the stimulus is seen to move behind a stationary window and fixation point whereas during pursuit, the window, fixation point and dot stimulus appear at roughly the same distance from the observer. However, we do not believe these are mitigating factors for two reasons. First, we show below that the change in perceived slant can be quantified by the change in perceived speed. One data set predicts the other and so there is no need to include additional factors such as differences in perceived distance or edge effects to explain the results. Second, we show in the corollary to Experiment 3 that the change in perceived slant disappears for higher values of shear, even though differences in perceived distance and edge effects remain.

A weakness of the slant-matching procedure is the inability to tell whether observers matched the 3D slant

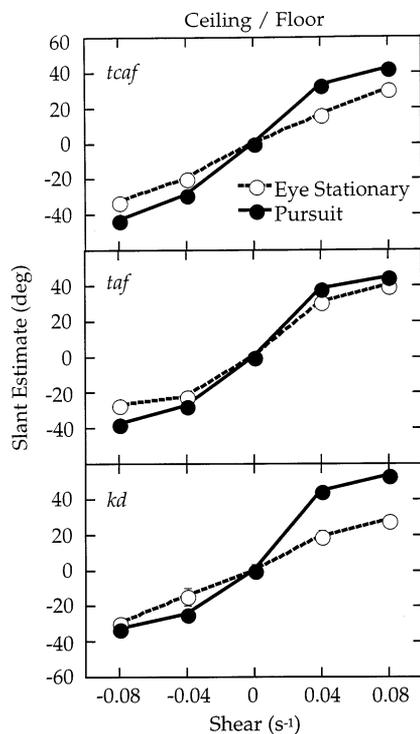


Fig. 6. Slant estimates of pursued (closed symbols) and eye-stationary (open symbols) stimuli in the ceiling/floor conditions. Different panels show results for different observers. The empirical functions have been forced to pass through the origin by subtracting the small bias found in the no-shear condition from each mean. For each observer, pursued stimuli produced the greater slant estimate. Error bars are ± 1 SE and are typically smaller than the symbol size.

of the stimuli or some other 2D property such as the perceived shear in the two intervals. This problem in interpreting the results of depth-matching is well-documented (e.g. Braunstein, 1994). Anecdotal evidence suggested that the matches were based on a 3D percept because all observers reported that both pursuit and eye-stationary stimuli appeared slanted for the larger values of shear. However, this does not rule out the possibility that the psychophysical judgement was based on matches made at some 2D level of representation. In the next experiment, a different technique was used to examine the hypothesised increase in slant, one that attempted to probe the 3D percept more directly.

5. Experiment 3: slant estimates for pursuit and eye-stationary stimuli

The perceived slant of pursuit and eye-stationary stimuli was examined using a slant-estimation technique. Observers viewed shearing, translating stimuli with and without pursuit and then estimated perceived slant using an adjustable ‘paddle’ presented on the display screen.

5.1. Procedure

The test stimuli were identical to those used in Experiment 2. The paddle consisted of a regularly-textured dot plane displayed in sequence on the same monitor as the moving test stimuli. When fronto-parallel to the observer, the paddle consisted of an evenly-spaced grid of dots subtending $12 \times 12^\circ$ with a dot density of ~ 1.56 dots/deg². The slant of the paddle could be adjusted in real time using movements of a computer mouse. This was achieved by pre-computing perspective projections of the paddle at 1° slant intervals about either a horizontal or vertical axis (depending on tilt direction) and then using the mouse to access the ‘frames’ of this ‘movie’. The paddle did not translate from side to side (or up and down in the case of right-edge/left-edge conditions) but remained in the centre of the screen.

Observers alternated between viewing the random-dot pattern and adjusting the slant of the paddle until they were satisfied with their slant estimate. To avoid stereotyped mouse movements, the gain of the mouse was randomised from setting to setting. The slant of the paddle was randomly selected at the beginning of each replication. No limit was placed on the number of reviews of the test stimuli or the amount of time the paddle was inspected. Each session comprised five replications of five shears (± 0.08 , ± 0.04 and 0 s⁻¹), presented in randomised blocks and combined with a translation speed of 4°/s. Pursuit and eye-stationary conditions were carried out in separate sessions. As in the first two experiments, the translation alternated leftward then rightward (or upward then downward) from view to view to minimise motion aftereffects.

5.2. Results and conclusions

Fig. 6 plots the mean slant estimates against the shear magnitude for the ceiling/floor conditions. Closed symbols correspond to the pursuit data and open symbols the eye-stationary data. Negative values of shear correspond to ‘ceilings’. Each panel shows data for a different observer. Fig. 7 plots the data for the right-edge/left-edge conditions. Negative values of shear correspond to ‘right edges’ in this case. For both data sets, small biases in slant estimates found in the no-shear condition were removed by translating each function to the origin. The mean bias for the pursuit condition was -1.99° (SE = 1.80°); for the eye-stationary condition, it was -1.15° (SE = 1.62°).

The empirical functions are sigmoidal in shape, as would be expected from the arctan function in Eq. (1) (see also Meese et al., 1995; Freeman et al., 1996). More importantly, slant estimates were always higher when the stimuli were pursued. Taken together, the results of Experiments 2 and 3 provide good evidence that perceived slant increases during pursuit.

Previous studies have shown perceived slant from motion cues to be consistently underestimated by observers (e.g. Braunstein, 1968; Harris et al., 1992). Indeed, underestimation is a recurring theme in many other judgements of motion-based depth (e.g. Rogers & Graham, 1979; Braunstein & Tittle, 1988). One explanation is that presenting a single, informative cue competes with the many other cues indicating that the display is flat, such as the lack of texture gradient (Braunstein, 1968). The resulting judgement may therefore be a compromise between multiple interpretations of depth. Providing more than one informative cue increases the accuracy of slant estimates (Braunstein,

1968), in some cases producing estimates that are veridical (Harris et al., 1992). Motion-based depth perception is therefore dependent on other cues present in the stimuli. This raises the possibility that a change in cue competition between the two eye-movement conditions may explain the present finding. For instance, competing cues-to-flatness may be more salient when the eye is stationary and so produce the effect found.

To investigate this issue, we adopted the model proposed by Freeman et al. (1996) for stimuli such as these. They modelled their slant estimation data using a formula similar to the following:

$$\hat{\sigma} = m \arctan\left(k \frac{S}{T}\right) \quad (2)$$

where $\hat{\sigma}$ is the slant estimate in Figs. 6 and 7. The parameter m sets the asymptotes of the arctan function, in other words its ‘amplitude’, whereas the parameter k scales its ‘slope’. Any difference in perceived slant due to the mismatch between retinal and extra-retinal estimates of object speed is captured by parameter k . (This parameter may also reflect differences in the estimation of shear, a point that is considered in detail later). The parameter m represents the degree of underestimation, possibly arising from competing cues-to-flatness. Finding large differences in fitted values of parameter m between the two eye-movement conditions would be difficult to account for within the current framework.

This issue was investigated by extending the range of shears in a further experiment. If m is similar in the two eye-movement conditions, then slant estimates in pursuit and eye-stationary conditions should eventually saturate at similar values. Put another way, the resulting arctan functions in the pursuit and eye-stationary conditions should have similar amplitudes but different slopes. Fig. 8 shows slant-estimation data for the ceiling/floor conditions over this extended range of shears. As is clearly seen, the slant estimates for pursuit and eye-stationary conditions saturate at similar points. The two empirical functions have similar amplitude but different slope, which suggests that the difference between the two conditions is not due to differences in the degree of underestimation. Fits of Eq. (2) to the data support this conclusion and are shown as lines in Fig. 8. The fitted values of m were similar (TCAF, pursuit = 0.58, eye-stationary = 0.70; TAF, pursuit = 0.67, eye-stationary = 0.80) and close to the values reported previously by Freeman et al. (1996). The values of k , however, were not similar (TCAF, pursuit = 1.33, eye-stationary = 0.48; TAF, pursuit = 1.24, eye-stationary = 0.53). The data and model fits strongly suggest that the change in perceived slant is due to the mismatch in perceived speed.

The data of Fig. 8 also demonstrate that differences in perceived slant occur for a restricted range of shears only. This is an inevitable consequence of the saturat-

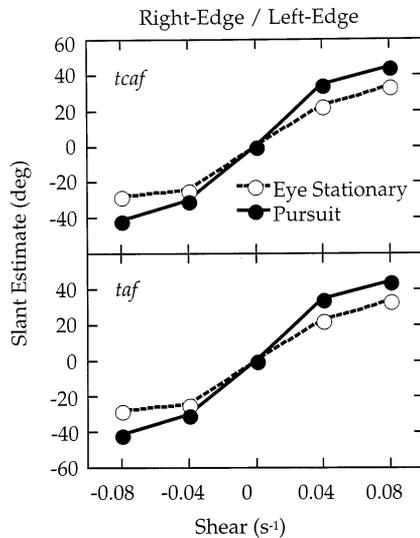


Fig. 7. Slant estimates for the right-edge/left-edge conditions. Symbol conventions and other details are the same as Fig. 6.

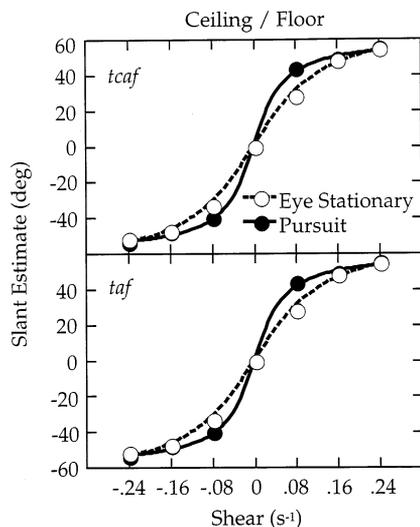


Fig. 8. Slant estimates saturate at the same value for large magnitudes of shear. The data are for the ceiling/floor condition. The lines represent fits of Eq. (2) to the data. Symbol conventions and other details are the same as Fig. 6.

tion effect. One interpretation is that increasing the magnitude of shear eventually swamps the difference between extra-retinal and retinal estimates of object speed. However, this interpretation assumes that estimates of shear are identical in the pursuit and eye-stationary conditions. Indeed, the discussion so far has concentrated exclusively on differences in extra-retinal and retinal estimates of translation speed. Inspection of Eq. (1) shows that shear must also be estimated by the visual system, which leads us to question whether changes in shear estimation between the two eye-movement conditions could contribute to the change in perceived slant. This is certainly possible because the retina is stimulated by shear alone in the pursuit condition. In contrast, the shear is embedded in a component of translation in the eye-stationary condition. Moreover, Nakayama (1981) has shown that thresholds for detecting shear increase when accompanied by a component of retinal translation, perhaps suggesting that shear estimates differ between the two eye-movement conditions. Of course, performance in detection experiments is limited by noise as well as signal, so it is difficult to extrapolate Nakayama's finding to the present work. For the same reason, it is also difficult to interpret the finding of Wertheim and Niessen (1986, cited in Wertheim, 1994) within the current context. They showed, in contrast to Nakayama, that the threshold for detecting the relative motion between two distinct objects is not affected by eye pursuit, despite the fact that the pursuit imposed additional retinal motion on their stimuli.

It remains unclear whether the increase in perceived slant during pursuit results from differences in perceived speed, perceived shear, or both. To examine this issue, we ask whether the change in perceived slant can be quantified solely by the change in perceived speed. Failure to predict the slant-estimation data from the perceived-speed data under the assumption that shear estimation is identical in the two eye-movement conditions would point to a more complicated explanation of our results.

6. Model

In order to predict the change in perceived slant using the perceived-speed data, we first need to assume the form of the relationship between input and output speed for retinal and extra-retinal processes. Freeman and Banks (1998) used a linear assumption to model the perception of head-centric object velocity (such as the Aubert–Fleischl phenomenon investigated in Experiment 1) and Freeman (1999) has used this assumption to predict performance in experiments on the Filehne illusion and the perceived direction of self-motion during eye pursuit. While we do not put the linear

assumptions to any strong test here, these previous studies suggest that it adequately describes many aspects of head-centric motion perception.

One can determine the ratio of extra-retinal to retinal gains (e/r) for the Aubert–Fleischl phenomenon on the basis of the linear assumption (Freeman & Banks, 1998). The individual gains e and r summarise the linear relationship between perceived extra-retinal speed and pursuit speed ($\hat{T}_p = eT_p$) and perceived retinal speed and actual retinal speed ($\hat{T}_{es} = rT_{es}$), respectively. The subscripts p and es denote 'pursuit' and 'eye-stationary' intervals, respectively. The gain ratio can be estimated from perceived-speed data by:

$$\frac{e}{r} = \frac{T_{es}}{T_p} \quad (3)$$

where T_{es} and T_p are the respective speeds of the eye-stationary and pursuit intervals at the perceived-speed match. In fact, this ratio was used to define relative perceived speed in Fig. 2.

The gain ratio can also be estimated from the slant-estimation data². Based on Eq. (1):

$$\hat{S}_p = \hat{T}_p \tan\left(\frac{\hat{\sigma}_p}{m}\right) \quad (4)$$

where \hat{S}_p is the visual system's estimate of the shear in the pursuit condition and $\hat{\sigma}_p$ is the slant estimate (in radians) during pursuit. The parameter m is the same as discussed above: it scales the amplitude of the arctan function in Eqs. (1) and (2).

A similar formula can be generated for the eye-stationary interval:

$$\hat{S}_{es} = \hat{T}_{es} \tan\left(\frac{\hat{\sigma}_{es}}{m}\right) \quad (5)$$

Note that we have assumed the same scaling parameter (m) for pursuit and eye stationary conditions. This would seem reasonable given the fits of Eq. (2) to the data of Fig. 8. For simplicity, m was set to the mean of the fitted values found, namely 0.69.

We can now test whether the shear estimates are the same in the two conditions by setting $\hat{S}_{es} = \hat{S}_p$. Combining this assumption with Eqs. (4) and (5):

$$\tan\left(\frac{\hat{\sigma}_{es}}{m}\right) = \frac{\hat{T}_p}{\hat{T}_{es}} \tan\left(\frac{\hat{\sigma}_p}{m}\right) \quad (6)$$

Using the linear relationships for extra-retinal and retinal estimates of translation speed and noting that translation speed was the same in the two eye movement conditions (i.e. $T_p = T_{es}$):

² One can construct models for the slant-matching data of Experiment 2 based on the same principles. However, the predictions for the slant-estimation data of Experiment 3 provide a better test of the ideas discussed, as slant estimates from more than one value of shear were obtained.

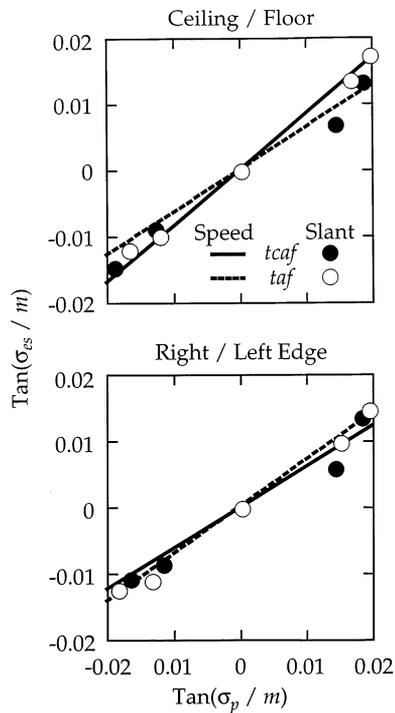


Fig. 9. Predicting perceived-slant data from perceived-speed data. The top panel shows the result of the analysis for the ceiling/floor conditions; the bottom panel, the result for right-edge/left-edge conditions. Closed symbols and solid lines are derived from the data of TCAF and open symbols and broken lines are derived from the data of TAF. The lines have a slope equal to the estimate of gain ratio (e/r) from the perceived-speed data of Fig. 3. Gain ratio can also be estimated from the perceived-slant data of Figs. 6 and 7 by plotting the tangent quantities defined by the axes. These are shown as symbols. The estimate of gain ratio from perceived-speed and perceived-slant experiments are in good agreement. The model assumes that the shear estimates made in the two eye-movement conditions are equal, and so we find no evidence that changes in perceived shear contribute to the change in perceived slant.

$$\tan\left(\frac{\hat{\sigma}_{es}}{m}\right) = \frac{e}{r} \tan\left(\frac{\hat{\sigma}_p}{m}\right) \quad (7)$$

Thus, when the two tangent quantities defined in Eq. (7) are plotted against one another, they should fall on a straight line with a slope equal to the gain ratio. That is, under the assumption that the shear estimates are equal in the two eye-movement conditions, we can estimate the gain ratio from the slant-estimation data. The symbols in Fig. 9 plot this relationship for the two authors, based on the data of Figs. 6 and 7 (their data was used because they were the only observers who participated in both the speed-matching and slant-estimation experiments). Different symbols are used to distinguish the two observers. The top panel corresponds to the ceiling/floor conditions whereas the bottom panel the right-edge/left-edge conditions.

Also plotted in the Figure are the expected slopes of these functions (i.e. the gain ratio) estimated from the speed-matching data using Eq. (3). These are shown as lines. The two estimates of gain ratio — that from the

speed-matching data (lines) and that from the slant-estimation data (symbols) — are in close agreement. For the experimental conditions studied, therefore, there is no evidence that shear is estimated differently in the two eye-movement conditions. The perceived-speed data quantitatively predicts the slant-estimation data. Perceived slant increases during pursuit not because other cues compete differentially in the two eye-movement conditions, nor because shear estimates change, but because perceived speed decreases.

7. Experiment 4: perceived slant decreases as pursuit speed rises

Another way to examine putative changes in shear estimation is to prevent them. We did this by presenting the same retinal shear at different pursuit speeds and recording slant estimates. So long as the eye movements are accurate, the retina is stimulated by the same magnitude of shear at each pursuit speed and so, presumably, shear estimates remain constant too. According to Eq. (1), however, we expected perceived slant to decrease because the available extra-retinal estimate of translation speed increases with pursuit speed.

7.1. Procedure

Pursuit stimuli were presented at four different translation speeds (1.25, 2.5, 5 and 10°/s) and two magnitudes of shear ($\pm 0.04 \text{ s}^{-1}$). Only ceiling/floor tilt directions were investigated; the translation was therefore horizontal. Each session comprised three replications of 16 conditions (two shears \times four speeds \times leftward and rightward pursuit). These were presented in a random order. Slant-estimation data were collapsed across eye movement direction before means were taken. In analysing eye movements, the recordings from the final two stimulus reviews that accompanied each slant estimate were used.

7.2. Results and conclusions

The data from the three observers (TCAF, TAF and KD) were similar, so Fig. 10 plots the mean slant estimates across observers as a function of pursuit speed. Positive slant estimates correspond to ‘floors’ and negative estimates to ‘ceilings’. The change in perceived slant is not great for lower speeds; as speed increases, however, perceived slant decreases more rapidly. The non-linear behaviour is a result of the arctan function in Eqs. (1) and (2). To demonstrate this Eq. (2) was fit to the data using a least-squares technique. The result is shown by the lines in the Figure. The model predicts the data well. Although we have

two free parameters and only four points per fit, we were unable to obtain good fits using either m or k in one-parameter versions of Eq. (2). The two-parameter model therefore appears to be a sensible description of these data. It explains why octave increases at slower speeds do not significantly alter perceived slant whereas the same logarithmic steps at higher speeds produce large changes.

Fig. 11 shows the eye movement data. Pursuit gains were somewhat variable at the lower speeds; as speed increased, however, pursuit gains stabilised. Moreover, they were accurate at the higher speeds. The eye-movement data suggest that at higher speeds, the retina was stimulated with approximately the same magnitude of shear. Presumably, then, the shear estimate was approximately the same at these speeds. We conclude that the change in perceived slant is due to the change in the size of the available extra-retinal, eye-velocity signal.

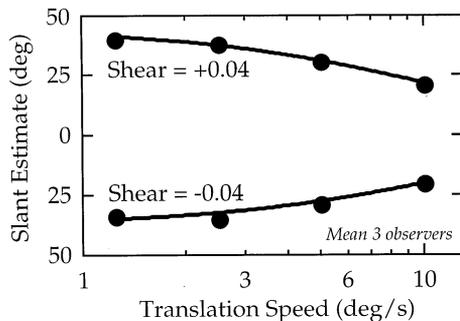


Fig. 10. Perceived slant decreases as pursuit speed rises. The data are the means across three observers for a shear of $+0.04$ and -0.04 s^{-1} conditions (top and bottom slant estimates, respectively). Assuming accurate pursuit, each translation speed was accompanied by an equal magnitude of shear for the two shears studied. Thus, the decrease in perceived slant is due to an increasing extra-retinal signal. The non-linear nature of the slant estimates is predicted by the arctan relationship between slant and the shear to translation ratio. This is seen by the fits of Eq. (2) to the data, depicted by the solid lines. Error bars are ± 1 SE, and are smaller than the symbol size.

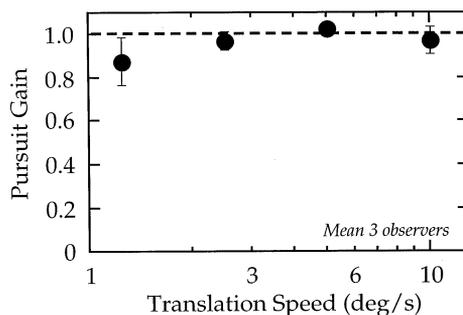


Fig. 11. Eye movement accuracy for the conditions of Experiment 4. Mean pursuit gain across the three observers is shown for each translation speed, collapsed across the two shear conditions. At the lower translation speed, pursuit eye movements were somewhat variable; however, as pursuit speed rises, pursuit accuracy stabilises. Error bars are ± 1 SE.

8. Discussion

The central question posed in this paper was to what extent pursuit eye movements affect the perception of depth from motion cues. As pointed out, eye movements alter the pattern of motion stimulating the retina; in many situations, the visual system must compensate for the change in order to extract a meaningful description of the scene. The perceived slant of surfaces moving at right angles to the observer was investigated, in part because the kinematics of this situation are simple to formulate. To a close approximation, the retina is stimulated by relative motion alone during accurate pursuit. When the eye is stationary, the stimulation consists of both relative motion and a component of translation. Because relative motion must be scaled by translation to obtain slant, the visual system needs to use an extra-retinal estimate of speed during eye pursuit. This raised the possibility that perceived slant should change in a predictable manner when pursuit and eye-stationary slant estimates are compared, because extra-retinal estimates of translation speed are typically lower than corresponding retinal estimates. Having demonstrated the Aubert–Fleischl phenomenon for slanted surfaces, perceived slant was then shown to increase during pursuit. The increase was predicted by the inverse relationship between translation speed and slant. Subsequent modelling indicated that the increase was entirely due to changes in perceived speed; there was no need to include changes in shear estimation between the two eye-movement conditions, nor differences in the saliency of conflicting cues-to-flatness. A final experiment showed that the perceived slant of a fixed retinal shear decreased as pursuit speed increased. This is to be expected because the available extra-retinal speed estimate increases with pursuit speed.

We have assumed throughout that motion-based slant perception involves separate estimates of relative motion (in our case, shear) and translation speed. This, of course, need not be the case when the eye is stationary. Eq. (1) is simply one way to describe the information present on the retina: proposing that the visual system estimates slant by combining separate estimates of shear and translation does not necessarily follow. For the case of slant perception during pursuit, however, there is more reason to suspect that separate estimates of shear and translation are combined. The relative-motion component can only produce a consistent estimate of slant if scaled by translation speed. But during pursuit, relative motion is provided retinally and translation extra-retinally. Thus, estimates of shear and translation are naturally separated into different processing streams. On the basis of the evidence we have presented, it is logical to conclude that they are then combined at some later stage to produce an estimate of

slant. If the visual system combines separate estimates of these components during pursuit, it is parsimonious to suggest that it does so when the eye is stationary.

Our analysis also assumes that shearing, translating stimuli are perceived as a rigid entity. Anecdotally, we found some of the high shear conditions in the saturation experiment of Fig. 8 yielded non-rigid percepts. The stimulus appeared to twist about the line of sight. Despite this, both observers were able to make consistent judgements of slant. The variability of slant estimates did not increase with the magnitude of shear, as is evident from Fig. 8. Moreover, there were no differences in the impression of rigidity between the two eye-movement conditions even when the shear was high, although this issue was not explicitly addressed by our experiments. In the earlier experiments, no observer reported non-rigid percepts.

The change in perceived slant suggests that metric estimates of depth from motion cues are unreliable. What implications might this have when more than one (informative) depth cue is present? Surface structure is defined by a variety of cues, some of which are not physically affected by eye movements. Texture gradients are one example. For the conditions of our experiments, more reliable cues such as texture may be given greater weight in the computation that yields a metric estimate of depth for any particular image region. This hypothesis is not, however, supported by the work of Braunstein (1968), who showed motion-based cues to slant were given a greater weight than texture-based cues in his cue-conflict experiments. But Braunstein did not compare of pursuit and eye-stationary conditions (nor were eye movements controlled in any way). It would therefore seem important to compare slant perception from texture cues under the same conditions of our experiments and then compare slant perception when both texture and motion cues are combined. If texture-based slant perception is unaffected by pursuit eye movements, then it is possible that perceived slant is less affected (if at all) when motion and texture cues are combined. We are currently exploring this idea.

Information about 3D structure involves computing ordinal properties of the scene, such as tilt, as well as metric quantities such as slant. Previous work on the interplay between extra-retinal signals and depth perception has concerned itself principally with perceived depth order and, moreover, has concentrated on extra-retinal signals for head movements (Hayashibe, 1991; Rogers & Rogers, 1992; Cornilleau-Peres & Droulez, 1994; Dijkstra, Cornilleau-Peres, Gielen & Droulez, 1995). Two conditions are typically compared: an 'active' condition, in which displays are perturbed according to the target depth modulation and the motion of the observer's head; and a 'passive' condition, in which the motion between target and observer is simulated in the display. These conditions parallel our eye-move-

ment conditions in many respects, in particular the notion that during active head movements, the retinal motion consists of relative-motion components alone. Using 1D corrugations in depth, Rogers and Graham (1979) showed that observers are able to report the correct shape of the target corrugations (see also Hayashibe, 1991). This would seem to imply the use of extra-retinal, head-velocity signals in the active condition because displays in which the relative-motion component is presented without extra-retinal or retinal translation information appear ambiguous: depth is perceived, but the phase of the corrugation is seen either correctly or 180° out-of-phase (Rogers & Rogers, 1992). Interestingly, many of these studies fail to control for eye movements made in active and passive conditions. Thus, it remains an open question whether the unambiguous depth order seen in the active condition is due to the additional vestibular cues associated with head translation or whether it is the inclusion of both vestibular and eye-movement signals that is responsible. We make this point because presumably in the active condition, the eye counter-rotates to maintain fixation on some element of the display; and in the passive condition, the eyes may move to stabilise the image. Thus, whether it makes sense to refer to these conditions as passive and active is debatable. Clearly, further work is needed to understand the interactions between extra-retinal and retinal motion cues to depth in experiments like these.

Intriguingly, Rogers and Graham (1979) and Ono and Steinbach (1990) report that the perceived depth modulation of 1D corrugations increases in the active condition. More recently, Yajima, Ujike and Uchikawa (1998) have shown analogous increases in perceived depth for head motions towards and away from depth modulated stimuli. Little explanation has been offered for these effects. Assuming that the passive condition in these papers did not involve any significant eye movement, then the theory tested in the present paper offers a possible explanation: the increase in perceived depth during head translation is the consequence of an analogous Aubert–Fleischl phenomenon for head movement. One way to test this is to investigate whether head-tracked objects appear to move more slowly than the same object viewed with no head movement. Of course, eye movements need to be appropriately controlled for. From Rogers and Grahams' description of their stimuli, eye movements in their experiments were not; however, a fixation point was provided by Ono and Steinbach.

In conclusion, we find that perceived slant increases when surfaces are tracked with an accurate eye movement. The increase can be quantified by the well-known mismatch between extra-retinal and retinal estimates of object speed that give rise to the Aubert–Fleischl phenomenon.

Acknowledgements

We would like to thank Jim Crowell and Mike Harris for providing comments on an earlier version of the manuscript and Casper Erkelens and Alexander Wertheim for comments during review. Some the experiments were reported in preliminary form at the annual meeting of the Association for Research in Vision and Ophthalmology (Freeman & Fowler, 1999). The work was supported by grants from the Royal Society and EPSRC (GR/N17638).

References

- Braunstein, M. L. (1968). Motion and texture as sources of slant information. *Journal of Experimental Psychology*, *78*, 247–253.
- Braunstein, M. L. (1994). Structure from motion. In A. T. Smith, & R. J. Snowden, *Visual detection of motion*. London: Academic Press.
- Braunstein, M. L., & Andersen, G. J. (1981). Velocity gradients and relative depth perception. *Perception & Psychophysics*, *29*, 145–155.
- Braunstein, M. L., & Tittle, J. S. (1988). The observer-relative velocity field as the basis for effective motion parallax. *Journal of Experimental Psychology: Human Perception and Performance*, *14*, 582–590.
- Bruss, A. R., & Horn, B. K. P. (1983). Passive navigation. *Computer Vision, Graphics and Image Processing*, *21*, 3–20.
- Cornilleau-Peres, V., & Droulez, J. (1994). The visual perception of three-dimensional shape from self-motion and object-motion. *Vision Research*, *34*, 2331–2336.
- Dichgans, J., & Brandt, Th. (1972). Visual-vestibular interaction and motion perception. In J. Dichgans, & E. Bizzi, *Cerebral control of eye movements*. *Bibliotheca ophthalmologica*, vol. 82. Karger.
- Dijkstra, T. M. H., Cornilleau-Peres, 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*, 453–462.
- Domini, F., & Caudek, C. (1999). Perceiving surface slant from deformation of optic flow. *Journal of Experimental Psychology: Human Perception and Performance*, *25*, 426–444.
- Freeman, T. C. A. (1999). Path perception and Filehne illusion compared: model and data. *Vision Research*, *39*, 2659–2667.
- Freeman, T. C. A., & Banks, M. S. (1998). Perceived head-centric speed is affected by both extra-retinal and retinal errors. *Vision Research*, *38*, 941–945.
- Freeman, T. C. A., & Fowler, T. A. (1999). Perceived slant increases during smooth eye pursuit. *Investigative Ophthalmology and Visual Science*, *40*, S167.
- 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.
- Georgeson, M. A., Freeman, T. C. A., & Scott-Samuel, N. E. (1996). Sub-pixel accuracy: psychophysical validation of an algorithm for fine positioning and movement of dots on visual displays. *Vision Research*, *36*, 605–612.
- Gibson, E. J., Gibson, J. J., Smith, O. W., & Flock, H. (1959). Motion parallax as a determinant of perceived depth. *Journal of Experimental Psychology*, *58*, 40–51.
- Harris, M. G. (1994). Optic and retinal flow. In A. T. Smith, & R. J. Snowden, *Visual detection of motion*. London: Academic Press.
- Harris, M. G., Freeman, T. C. A., & Hughes, J. (1992). Retinal speed gradients and the perception of surface slant. *Vision Research*, *32*, 587–590.
- Hayashibe, K. (1991). Reversals of visual depth caused by motion parallax. *Perception*, *20*, 17–28.
- von Holst, E. (1954). Relations between the central nervous system and the peripheral organs. *British Journal of Animal Behaviour*, *2*, 89–94.
- Koenderink, J. J. (1986). Optic flow. *Vision Research*, *26*, 161–180.
- 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.
- Longuet-Higgins, H. C., & Prazdny, K. (1980). The interpretation of a moving retinal image. *Proceedings of the Royal Society of London*, *B208*, 385–397.
- Mack, A., & Herman, E. (1973). Position constancy during pursuit eye movement: an investigation of the Filehne illusion. *Quarterly Journal of Experimental Psychology*, *25*, 71–84.
- Mack, A., & Herman, E. (1978). The loss of position constancy during pursuit eye movements. *Vision Research*, *18*, 55–62.
- Maxwell, A. E. (1959). Maximum likelihood estimates of item parameters using the logistic function. *Psychometrika*, *24*, 221–227.
- Meese, T. S., & Harris, M. G. (1997). Computation of surfaces slant from optic flow: orthogonal components of speed gradient can be combined. *Vision Research*, *37*, 2369–2379.
- Meese, T. S., Harris, M. G., & Freeman, T. C. A. (1995). Speed gradients and the perception of surface slant: analysis is two-dimensional not one-dimensional. *Vision Research*, *35*, 2879–2888.
- Nakayama, K. (1981). Differential motion hyperacuity under conditions of common image motion. *Vision Research*, *21*, 1475–1482.
- Ono, H., & Steinbach, M. J. (1990). Monocular stereopsis with and without head movements. *Perception & Psychophysics*, *48*, 179–187.
- Rogers, B. J., & Graham, M. (1979). Motion parallax as an independent cue for depth perception. *Perception*, *8*, 125–134.
- Rogers, S., & Rogers, B. J. (1992). Visual and nonvisual information disambiguate surfaces specified by motion parallax. *Perception & Psychophysics*, *52*, 446–452.
- Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology*, *45*, 205–217.
- Wertheim, A. H. (1994). Motion perception during self-motion-the direct versus inferential controversy revisited. *Behavioural and Brain Sciences*, *17*, 293–311.
- Wertheim, A. H., & Niessen, M. W. (1986). The perception of relative motion between objects during pursuit eye movements. *Perception*, *15*, 49.
- Wyatt, H. J. (1998). Detecting saccades with jerk. *Vision Research*, *38*, 2147–2153.
- Yajima, T., Ujike, H., & Uchikawa, K. (1998). Apparent depth with retinal image motion of expansion and contraction yoked to head movement. *Perception*, *27*, 937–949.