

Available online at www.sciencedirect.com

SCIENCE @ DIRECT®

Vision Research 44 (2004) 3025–3034

Vision
Researchwww.elsevier.com/locate/visres

Perceiving depth order during pursuit eye movement

Jenny J. Naji, Tom C.A. Freeman *

School of Psychology, Cardiff University, Tower Building, Park Place, CF10 3AT, Wales, UK

Received 9 February 2004; received in revised form 23 June 2004

Abstract

Pursuit eye movements alter retinal motion cues to depth. For instance, the sinusoidal retinal velocity profile produced by a translating, corrugated surface resembles a sinusoidal shear during pursuit. One way to recover the correct spatial phase of the corrugation's profile (i.e. which part is near and which part is far) is to combine estimates of shear with extra-retinal estimates of translation. In support of this hypothesis, we found the corrugation's spatial phase appeared ambiguous when retinal shear was viewed without translation, but unambiguous when translated and viewed with or without a pursuit eye movement. The eyes lagged the sinusoidal translation by a small but persistent amount, raising the possibility that retinal slip could serve as the disambiguating cue in the eye-moving condition. A yoked control was therefore performed in which measured horizontal slip was fed back into a fixated shearing stimulus on a trial-by-trial basis. The results showed that the corrugation's phase was only seen unambiguously during the real eye movement. This supports the idea that extra-retinal estimates of eye velocity can help disambiguate ordinal depth structure within moving retinal images.

© 2004 Elsevier Ltd. All rights reserved.

Keywords: Depth; Motion; Cues; Extra-retinal; Eye movement

1. Introduction

One way the visual system extracts depth information from moving images is by analysing the patterns of movement that play out across the retina. At any point in time, the spatial gradients of retinal velocity provide quite detailed information about the relative depths of points in the scene (Harris, 1994; Koenderink, 1986). The spatial structure of retinal motion is therefore a useful cue to depth, allowing the observer to recover properties such as the three-dimensional structure of an object's surface (Braunstein & Tittle, 1988; Domini & Caudek, 1999; Domini & Caudek, 2003; Freeman, Harris, & Meese, 1996; Gibson, Gibson, Smith, & Flock, 1959; Harris, Freeman, & Hughes, 1992; Meese &

Harris, 1997; Rogers & Graham, 1979; Wallach & O'Connell, 1953). However, motion also stimulates a variety of responses from the eye-movement system. In particular, the eyes tend to pursue a moving stimulus unless the observer is provided with a stationary fixation point. Pursuit eye movements introduce global components of retinal image motion that add vectorially to any retinal motion cue to depth. In normal free viewing, therefore, the stimulus for recovering depth from motion is quite different from that often portrayed in the literature.

There has been considerable debate over the way the visual system compensates for pursuit eye movements. Work has focussed on compensation during the perception of object velocity and self-motion (Freeman, 1999; Freeman & Banks, 1998; Lappe, Bremmer, & van den Berg, 1999; Royden, Banks, & Crowell, 1992; Turano & Massof, 2001; Wertheim, 1994). Much of this work suggests that observers use extra-retinal information about eye velocity as a means of interpreting the sensed

* Corresponding author. Tel.: +44 0 29 20 874554; fax: +44 2920 874858.

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

motion on the retina. Here we ask whether extra-retinal, eye-velocity signals also play a role in the judgement of motion-defined depth.

Fig. 1 depicts a sinusoidal depth corrugation moving horizontally at right angles to the observer. When the eye is stationary (Fig. 1A), nearer points move faster on the retina, with the particular corrugated shape revealed by the specific way in which velocity changes across the image. Fig. 1C shows the retinal motion pattern produced when the observer tracks the corrugation with a pursuit eye movement. Assuming the observer pursues accurately, the pattern now corresponds to a sinusoidal shear as shown. The spatial phase of the corrugation's depth profile, that is whether the surface appears 'top-far' or 'top-near', is determined by the combination of the relative motion component (shear in this case) and translation (Domini & Caudek, 1999; Freeman & Fowler, 2000; Freeman et al., 1996; Harris, 1994). In the absence of translation information the spatial phase is easily confused because the shearing pattern approximates the orthographic projection of a corrugation rotating about a vertical axis (Hayashibe, 1991; Rogers & Collett, 1989—see Fig. 1B). It is not an exact rendition because, for example, the texture does not compress horizontally over time (Liter & Braunstein, 1998). Nevertheless, the shearing pattern is reminiscent of the type of stimuli used to generate the kinetic depth effect (Wallach & O'Connell, 1953). In both cases the lack of perspective information makes it difficult to determine which part of a surface is more distant. The ambiguity could be resolved in the case shown in Fig. 1C if the observer knew the direction of travel.

In an earlier study we showed that extra-retinal, eye-velocity signals contribute to the judgement of depth amplitude. Specifically, we showed that the decrease in perceived translation speed that occurs during pursuit results in an increase in perceived slant (Freeman &

Fowler, 2000). The result was predicted from the fact that depth amplitude (e.g. slant) is determined by the ratio of shear to translation, a relationship that has been used, for example, to explain how depth-sensitivity changes with head-translation speed (Ujike & Ono, 2001). More recently, Nawrot (2003) showed that eye movement information also helps disambiguate depth order. To establish this he examined depth profiles perceived when shearing motion aftereffects were combined with a variety of head and eye movements. Here we ask whether extra-retinal, eye-velocity signals also contribute to the perception of depth order when real motion is pursued.

Observers were asked to judge the spatial phase of sinusoidal corrugations in the three conditions shown in Fig. 1. In the first condition shown on the left, retinal translation and shear were viewed with stationary fixation. Spatial phase is determined by the speed of motion on the retina, according to a faster-is-nearer rule. In the third condition showed on the right, observers pursued the stimulus. Applying the same heuristic to the resulting retinal motions would lead to ambiguous interpretations of spatial phase. To recover the correct spatial phase observers need to know how the stimulus is translating, which can be obtained from an extra-retinal, eye-velocity signal. Combining retinal shear with extra-retinal translation is equivalent to computing the head-centred velocity of each point in the image, though whether the visual system actually performs this calculation is beyond the scope of the current paper. If observers ignored the extra-retinal signal, however, their judgements would be based on retinal motion alone. Depth judgements would therefore resemble the ambiguous depth structure seen when the eye is stationary and the translation is removed (Fig. 1B).

Interpreting the results rests largely on the observer's ability to pursue accurately because failure to do so introduces retinal slip into the image which may help disambiguate spatial phase. On average, eye-movement recordings showed a small but persistent temporal phase lag with respect to the sinusoidal modulation used in the experiments. A yoked control was therefore performed to see whether the resulting retinal slip was sufficient to correctly judge spatial phase.

2. Experiment 1

Stimuli like those described by Fig. 1 were presented in a single-interval forced choice paradigm. Observers were forced to discriminate between the two possible spatial phases ('top-far' and 'top-near') for a range of shears. This allowed frequency-of-seeing curves to be constructed as a function of the relative motion direction and amplitude (e.g. Bradshaw & Rogers, 1999). Curves resembling typical psychometric functions would

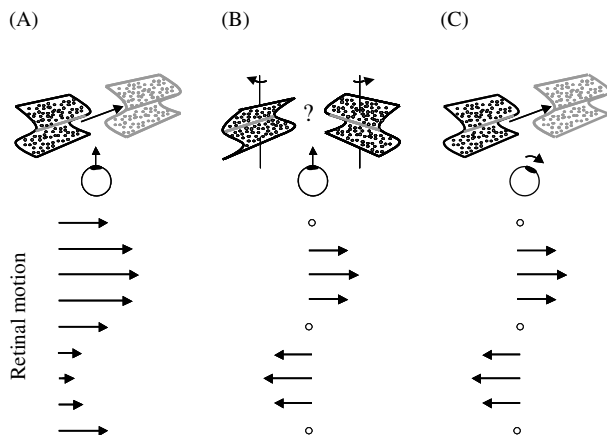


Fig. 1. Schematic of the three main conditions investigated: (A) eye stationary with translation; (B) eye stationary with no translation and (C) eye moving.

indicate stimuli whose spatial phase appeared unambiguous, such as predicted for situations in which shear was combined with translation regardless of eye movement (Fig. 2A and C, top). Conversely, any ambiguity in perceived spatial phase would lead to non-monotonic curves, with phase choices centred on an average frequency of 50% assuming that depth sign changed randomly from trial to trial. This is the type of function predicted for stimuli that do not translate (Fig. 2B, top).

This type of non-monotonic behaviour could also be produced by an inability to see depth when making the binary discrimination. A second forced choice was therefore included in which observers had to label stimuli as ‘three-dimensional’ or ‘flat’. The latter term encompassed those stimuli that appeared ‘two-dimensional and non-rigid’. To be defined as ambiguous, frequency-of-seeing curves needed to display not only a non-monotonic relationship between shear and perceived spatial phase (Fig. 2B, top) but also a peaked relationship between shear and perceived flatness (Fig. 2B, bottom).

2.1. Methods

2.1.1. Stimuli

Stimuli consisted of moving random-dot patterns displayed at 100Hz on the black background of a Mitsubishi Diamond Pro 20 monitor. This was driven by a VSG2/3 graphics board under PC control. Patterns had a dot density of 4 dots/deg² and were viewed through a square clipping window that was 10° wide and surrounded a central fixation point. Dots were also clipped from a circular region of radius 1° centred on the fixation point. The motion of the window was yoked to that of the fixation point so that in the eye-moving condition the fixation point and window moved in unison.

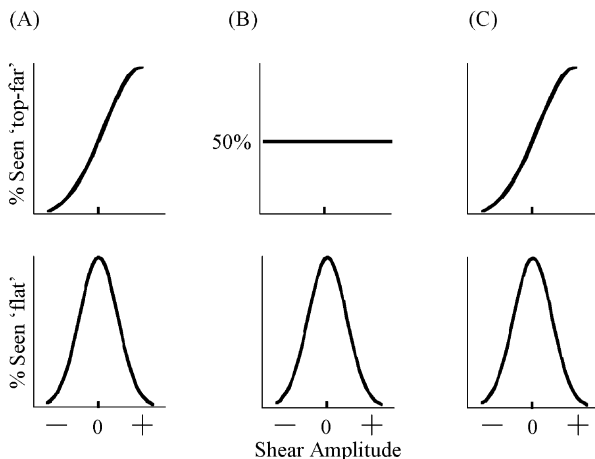


Fig. 2. Frequency-of-seeing curves assuming either retinal or extra-retinal estimates of translation help disambiguate depth order in shearing patterns. Top row corresponds to judgements of spatial phase, bottom row to judgements of three-dimensionality. Columns are in the condition order defined in Fig. 1.

Stimuli comprised two motion components, a sinusoidal shear and a horizontal translation. When combined these produced the sinusoidal velocity profile shown in Fig. 1A. The whole display was then modulated sinusoidally in time at a frequency f , which allowed observers a relatively long continuous view of the stimuli. The translation component therefore oscillated from side-to-side whilst preserving its temporal phase relationship with the shear. The horizontal component of velocity was defined as:

$$v_x = 2\pi f \cos(2\pi ft) \times [T + S \sin(2\pi f_s y)]$$

where T is the translation amplitude, S the shear amplitude, f_s the spatial frequency of the sinusoidal shear (equivalent to the spatial frequency of the depth corrugation) and y the vertical dot position with respect to screen centre. According to this relationship the translation component moved over a fixed distance regardless of the temporal modulation f . In all experiments, the spatial frequency f_s was fixed to 0.1 cpd, yielding one full period within the 10° window. This is reasonably close to the peak of the depth sensitivity curves reported previously for head movements (Hogervorst, Bradshaw, & Eagle, 2000; Rogers & Graham, 1982). The temporal modulation f was 0.5 Hz in the main experiments.

We defined negative shear as that producing dots above the fixation point moving initially right on the screen when no translation was present. Temporal phase was fixed, which meant the translation component had a fixed phase as well, moving first to the right. The combination of negative shear and translation therefore produced an oscillating corrugation with its upper peak nearer to the observer than the bottom, which we refer to as ‘top-near’ (see Fig. 1A).

In the eye-stationary-with-translation condition, the fixation point and square clipping window remained stationary and T was set to 1°. In the eye-stationary-no-translation condition, T was set to 0. In the eye-moving condition, T was set to 1° and the fixation point and clipping window moved with the same translation amplitude. The window and fixation point therefore moved over a distance of 2° as they oscillated back and forth in time with the dots, regardless of the temporal frequency used.

All stimuli were viewed monocularly at a distance of 57.3 cm. The experiments were conducted in a dark room to eliminate any external reference points. The head was stabilised in a chin-and-cheek rest.

2.1.2. Procedure

The three conditions were examined in separate experimental sessions. Each session investigated seven amplitudes of shear, including zero. These were shown in 10 randomised blocks, giving 70 trials per session in total. Each session took approximately 25 min to complete. Observers undertook three separate replications

per condition. The amplitudes of shear were tailored to the sensitivity to depth found in each of the conditions (see below). On each trial the fixation point appeared on its own for 2 s (i.e. for one period) followed by 4 s of temporally-modulated dot motion. For the eye-moving condition, this corresponded to a fixation point moving on its own for one period, followed by two periods of shear and translation. In the two eye-stationary conditions the fixation point was stationary throughout. Following each trial, observers first judged the phase of the corrugation ('top-far' or 'top-near') and then, immediately following this, judged whether the stimulus appeared 'flat' or 'three-dimensional'.

2.1.3. Eye-movement recording and analysis

Eye movements were recorded with a head-mounted video-based eye tracker (Applied Science Laboratories Series 4000). Eye-position recordings were made at a sampling rate of 50 Hz and analysed off-line using customised software written in Matlab. The initial part of each trial, consisting of 2 s of fixation point alone, was not analysed. The remainder of the recording was first low-pass filtered and then, for the purposes of detecting saccades, eye velocity computed by taking a time derivative. Saccades were identified using a velocity threshold region with width 20°/s above and below the target velocity profile (Ebisawa, Minamitani, Mori, & Takase, 1988). Trials containing saccades were discarded. Eye-movement accuracy was assessed by fitting sinusoids

to the position records using a least-squares technique, with amplitude, phase and DC as free parameters. This is equivalent to taking the Fourier transform of the position record and examining the amplitude and phase spectra at the fixation-target frequency (Collewijn & Tamminga, 1984). Pursuit gain was computed by dividing by the pursuit-target amplitude, T .

2.1.4. Observers

Five observers participated in the experiment. Three were naïve to the purposes of the study (JHS, BAN, CHT) and two were not (JJN, TCAF). All except BAN were experienced psychophysical observers and each had normal or corrected-to-normal vision.

2.2. Results and conclusions

2.2.1. Psychophysics

Fig. 3 plots the frequency-of-seeing curves for depth and flatness judgements. The layout is the same as Fig. 2. The open symbols show individual data and the thick lines with closed symbols the mean across observers. Spatial-phase judgements for the eye-stationary-with-translation condition resembled typical psychometric functions (left top). They were also coupled with flatness judgements that peaked at 0 shear (left bottom). When the translation was removed, the spatial-phase judgements became non-monotonic (middle top), though flatness judgements still peaked at 0 shear (middle bottom).

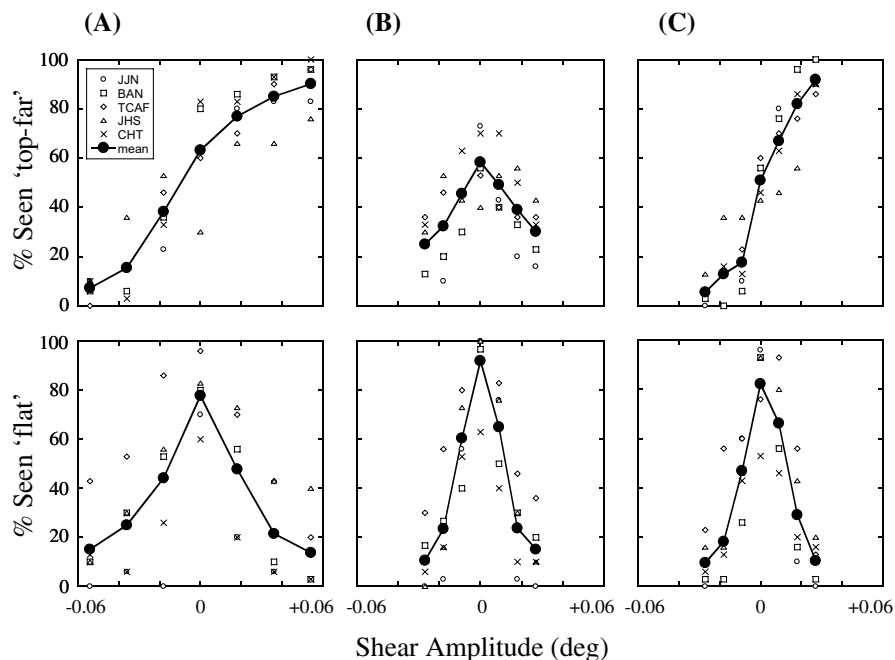


Fig. 3. Judgements of spatial phase (top row) and three-dimensionality (bottom row) in same format as Fig. 2. Open symbols correspond to individual observer performance aggregated across sessions. Closed symbols correspond to means: (A) eye stationary with translation; (B) eye stationary with no translation and (C) eye moving.

The data suggest a bias towards ‘top-near’ in this condition. However, when retinal shear was viewed with an eye movement, judgement of spatial phase became monotonic once more (right top). To the extent that the eye movements were accurate (see below), these changes in the perception of spatial phase support the conclusion that observers used extra-retinal, eye-velocity information to interpret retinal shear.

Two other features of the psychophysical data are worth considering. First, the slopes of the phase-judgement curves correlate with the width of the flatness curves in the two translation conditions (compare left and right columns). This is unsurprising. The ability to report spatial phase consistently will be a function of the observer’s sensitivity to retinal shear—as shear approaches zero, observers are more likely to report a stimulus that appears flat. The second more-important feature concerns the change in slope of the phase-judgement functions (compare top-left and top-right panels). On average these are steeper when the eye pursued. To quantify this effect we determined phase-discrimination thresholds by fitting logistic functions to the individual data and computing the just-noticeable-difference (JND, Wetherill & Levitt, 1965). Fig. 4 shows the result. In three cases the difference between the two conditions was quite large; for the other two the difference was negligible. Hence there is some evidence that spatial-phase sensitivity was greater in the eye-moving condition. It may be worth noting that the two observers who did not show any great sensitivity difference also showed reasonably large horizontal shifts in the psychometric function for the eye-stationary-with-translation condition. Reasons why a change in slope might exist between these two conditions are taken up in the Section 4.

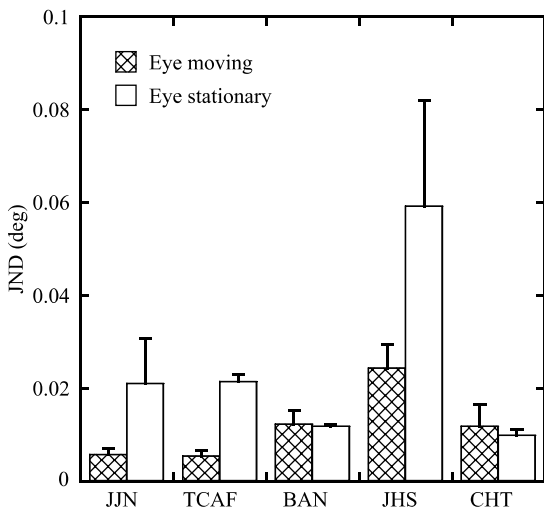


Fig. 4. Phase discrimination thresholds for the two translation conditions. Eye-moving thresholds are shown as hashed bars and eye-stationary thresholds as open bars. Error bars are ± 1 SE.

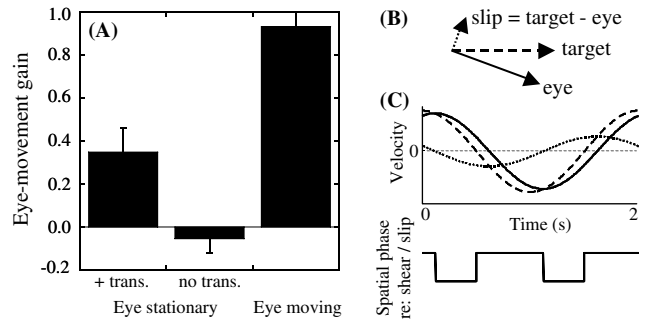


Fig. 5. (A) Mean pursuit gain across observers. Error bars are ± 1 SE; (B) computation of retinal slip, representing sinusoids as vectors. The vector labelled ‘eye’ is based on mean amplitude and phase of the recorded eye movements and (C) corresponding velocity sinusoids. The step function shows spatial phase according to the sign of the shear/slip ratio.

2.2.2. Eye movements

The eye-movement data make unequivocal interpretation of the psychophysical data quite difficult. Fig. 5A shows the mean eye-movement gains. The least serious problem was the low-amplitude tracking evident in the eye-stationary-with-translation condition (left bar). There was some variability across observers in this condition, with three of the five observers primarily responsible for the unexpectedly high gains found (TCAF and BAN had negligible gains of 0.1). The more serious problem accompanies the eye movements made in the other two conditions. The gain data suggests pursuit amplitude was quite accurate in the eye-moving condition and also fixation reasonably stationary in the eye-stationary-no-translation condition. However, a small and persistent lag in temporal phase accompanied the former. The lag had a mean temporal phase of 21.51° ($SE = 2.35^\circ$), which is equates to a delay of approximately 120ms. This degree of phase lag introduces appreciable horizontal retinal slip into the image during an eye movement (we did not analyse vertical components). To assess its impact, the three relevant sinusoidal velocities were treated as vectors in a 2D space, with length defining amplitude and direction defining temporal phase. Fig. 5B shows the vector representation, with the depicted eye-movement vector based on the mean amplitude and phase of pursuit found in the eye-moving condition. Slip was obtained by subtracting the pursuit vector from the fixation target vector. On average, the horizontal slip was about one-third of the fixation target and led by some 70° (~ 390 ms). These average velocity sinusoids are shown in Fig. 5C.

Closer inspection of the data showed there were few trials containing negligible horizontal slip. Retinal slip could therefore have acted as the disambiguating cue in the eye-moving condition. Importantly, the slip and shear were out-of-phase with one another, so whether

the retinal slip could be used to judge the spatial phase unambiguously is difficult to say. In theory, the corrugation's spatial phase is determined by the sign of the ratio between shear and translation. If the *temporal* phases of these two components are neither in-phase nor perfectly anti-phase, as was the case with the slip and the shear, the sign of the ratio flips back and forth. The step-function at the bottom of Fig. 5C shows how the sign of the ratio between shear and slip changed given the average eye movement found. Over the course of one period of stimulation, the ratio was positive for approximately 62% of the time. The question is whether this was enough for the spatial phase to be perceived unambiguously without using an extra-retinal signal.

A possible solution to this problem is to try and improve eye-movement accuracy by decreasing the temporal frequency of the sinusoidal modulation. Fig. 6 shows the results of a subsidiary experiment on two of the observers (one author and one naïve), in which the eye-moving condition was investigated over a range of pursuit-target frequencies. The eye-movement data in Fig. 6A and B shows that even at a relatively low frequency of 1/3 Hz, significant retinal slip remained. The phase-discrimination JNDs were similar to those found in the main experiment (Fig. 6C).

There is no pursuit-target motion that can guarantee all observers will pursue with perfect accuracy on each trial. Even when observers do not engage in a concurrent perceptual judgement, small errors in pursuit of sinusoidal targets remain (see Fig. 6 of Barnes, 1993; or Fig. 3.20 of Carpenter, 1988). The closest one can get to slip-free stimulation is to stabilise the image, though in the context of the current experiments viewing the pursuit-target would have to remain closed-loop as in the sophisticated experiments of Turano and Massof (2001). In doing so, however, little is learnt about normal unstabilised viewing. For this reason we designed the following yoked control to better investigate the influence of retinal slip in judging depth-from-motion during eye movement.

3. Experiment 2

Two conditions were compared. The first repeated the eye-moving condition of Experiment 1, this time with only one period (2s) of dot motion displayed on each trial. The second yoked-slip condition aimed to simulate as closely as possible the horizontal slip created during the eye movement made in the first condition, but this time with the eye stationary. Both conditions contained retinal shear and so both contained the same retinal information with which unambiguous depth judgements could be made. The vertical slip was ignored. The second condition was yoked to the first, in the sense that the retinal slip was based on trial-by-trial and sample-by-sample horizontal eye movements recorded in the first. Trial order was identical in the two conditions.

In Experiment 1 some observers were unable to inhibit pursuit in the eye-stationary-with-translation condition. This perhaps presents a problem when the translation is replaced by simulated slip because unwanted eye movements would reduce the similarity between the eye-moving and yoked-slip conditions. A careful trial-by-trial error analysis was therefore performed, details of which are given below.

3.1. Methods

The dot stimuli for the eye-moving condition were identical to those used in Experiment 1, with duration truncated to one period of modulation. Prior to this the fixation target appeared on its own for one period. It moved for one cycle before the dot pattern appeared in the eye-moving condition, or remained stationary at all times in the yoked-slip condition. Observers made spatial-phase and flatness judgements as before.

Recorded eye movements were used to determine the horizontal slip for each eye-moving trial on a sample-by-sample basis. This was computed offline in MatLab after each eye-movement session (i.e. after 70 trials). Position samples were first low-pass filtered and

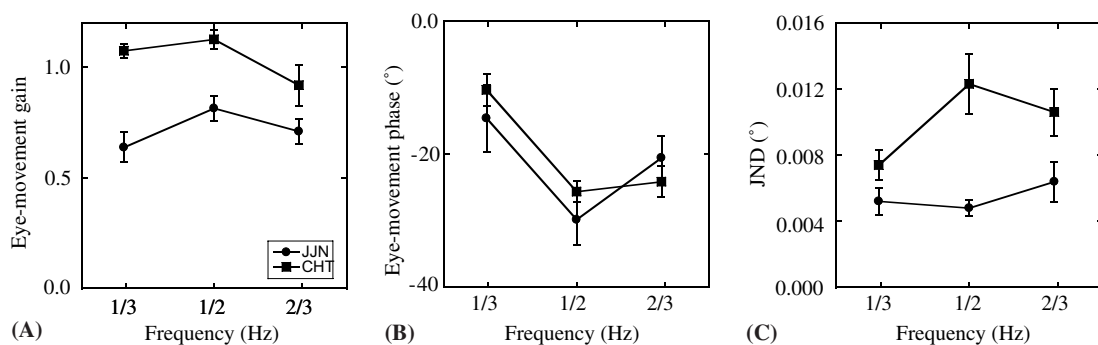


Fig. 6. Results for the eye-moving condition over a range of frequencies for two observers: (A) mean pursuit gain; (B) mean phase and (C) phase-discrimination thresholds. All error bars are ± 1 SE.

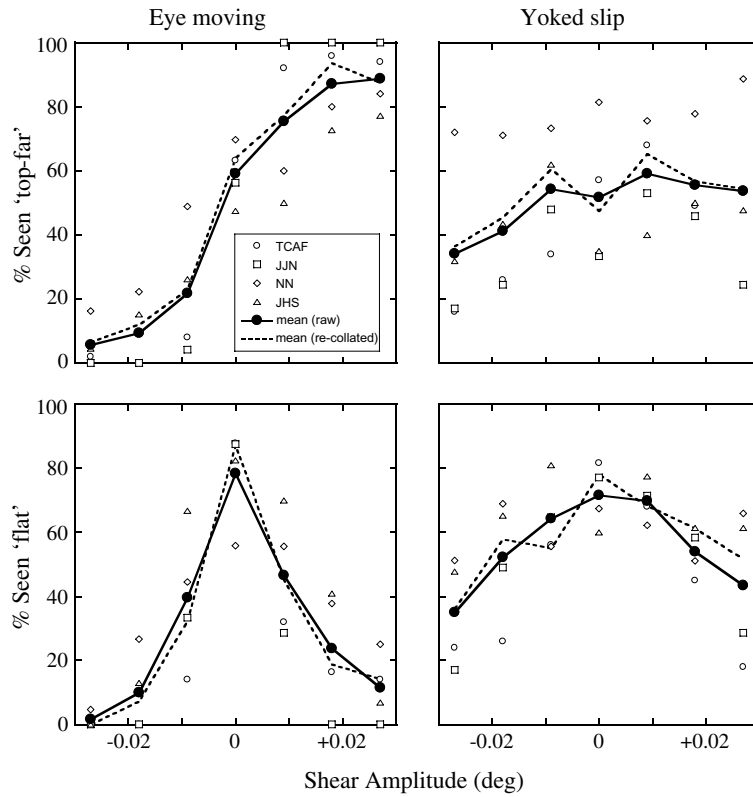


Fig. 7. Judgements of spatial phase (top) and three-dimensionality (bottom) for eye-moving and yoked-slip conditions. The latter simulated the slip in the former but without eye movement. Open symbols correspond to individual observer performance aggregated across sessions. Closed symbols and solid lines correspond to means of all trials except those containing saccades. Dotted lines are re-collated functions with poorly-correlated trial pairs removed.

the time derivative taken. Slip velocity was then determined by subtracting eye movements from the fixation-target sinusoid, using cubic-spline interpolation to resolve the fact that the eye tracker sampled at half the rate of the display. The result was stored to disk and then a session of the yoked-slip condition was run immediately, using the same trial order. This consisted of viewing shear plus horizontal slip but with the eye stationary. We therefore could not mimic the retinal slip associated with the fixation point itself and also decided not to perturb the viewing window either. For compatibility between conditions saccadic trials were not removed before running the yoked-slip condition. Hence a small percentage of yoked trials contained high-speed translation as the retinal effects of the saccade were simulated.

Observers carried out five sessions of each condition in alternating order, yielding 2×350 trials in total. Observer BAN was unable to participate.

3.2. Results and conclusions

Yoked-trial pairs containing saccades in either the eye-moving or eye-stationary condition were first ex-

cluded before frequency-of-seeing curves collated. Fig. 7 plots the result. Open symbols correspond to the depth and flatness judgements of individual observers and closed symbols the mean. In the eye-moving condition (top left) spatial-phase judgements were typically sigmoidal. The depth judgements were therefore similar to those found in Experiment 1. In the yoked-slip condition spatial-phase judgements were flat (top right). In both conditions, perceived-flatness curves peaked at 0 shear, though the function was not as sharp in the yoked-slip condition. The data suggest that the simulated slip could not be used to disambiguate spatial phase. It appears the eye movement was essential for judging spatial phase unambiguously.

It is possible the flattened depth judgments in the yoked-slip condition were the result of considerable reduction in depth sensitivity, perhaps produced by increased external noise arising from fixational jitter. Indeed, fixational jitter and also the inadvertent ocular-following reported in Experiment 1 would reduce the retinal similarity between the two conditions. The eye-movement recordings were therefore used to identify and remove any yoked-trial pairs in which actual and simulated slip were deemed dissimilar. To do this, the

100 samples per trial of actual slip and intended slip were plotted against each other and the trial pair removed if the square of the correlation coefficient¹ was less than 0.5, or the slope of the linear relationship was greater than $\pm 10\%$ of unity. Around 2/3 to 3/4 of trial-pairs were excluded in this manner. Despite this, the psychometric functions re-collated from the remaining trials closely resembled the originals. This can be seen in Fig. 7 by comparing the means of the original data (solid lines) to the means of re-collated data (dotted lines).

4. General discussion

The experiments described here suggest that extra-retinal signals are used to judge depth order from retinal motion cues during pursuit. The perceived spatial phase of a corrugated surface defined by a sinusoidal shear was investigated with and without eye movement. In the absence of translation, depth order was seen ambiguously. Stimuli appeared three-dimensional but with a spatial phase that varied from trial to trial. Conversely, when the shear was translated the spatial phase appeared unambiguous whether the stimulus was pursued or not. The data could not be explained on the basis of the horizontal translation information contained in the retinal slip. Though the eye exhibited a small but persistent phase lag at all frequencies studied, the phase-advanced slip could not be used to disambiguate depth order when it was simulated in an eye-stationary condition. This suggests that during pursuit a retinal estimate of relative motion such as shear is combined with an extra-retinal estimate of translation.

4.1. Signals and phases

In order to do so, the visual system may need to overcome temporal misalignments between the relevant signals. The eye lagged the fixation target, suggesting that extra-retinal translation signals and retinal shear signals might not have been completely in-phase. Of course, just because the physical phases are misaligned does not necessarily mean the underlying neural signals are as well. Moreover, it is unclear how large the temporal-phase difference should be before performance is seriously affected. It is therefore reassuring to note that in a study of eye-movement compensation, Freeman, Banks, and Crowell (2000) found only a small temporal-phase difference between retinal and extra-retinal velocity signals. If this remains true for signals encoding eye-movement velocity and retinal shear, then perhaps a temporal-phase difference on this scale is of little consequence.

¹ For $df = 98$, the one-tailed critical value of $r = 0.2$ ($r^2 = 0.04$) at $p = .05$.

4.2. Object rotations during pursuit

Experiment 1 showed that shear on its own appears three-dimensional but with ambiguous spatial phase. This is because the motion pattern approximates the orthographic projection of a sinusoidal corrugation rotating about a vertical axis, though to reiterate, it is not an exact rendition because the texture does not compress horizontally over time. The rotation could be seen quite clearly in our displays. Potentially the same rotating interpretation could be made in the eye-moving condition—indeed, some observers reported it when prompted, but its presence was somewhat ephemeral. This agrees with anecdotal reports by Rogers and Collett (1989). They found perceived rotation more prevalent when motion cues were placed in conflict with binocular cues, but virtually non-existent for a condition similar to the eye-moving condition used here. Regardless of whether rotation was seen or not, all our observers were biased to the faster-is-nearer rule. Thus they produced judgements of spatial phase suggesting they combined shear with an extra-retinal estimate of translation.

Nevertheless, objects can both rotate and translate regardless of how the eye moves and so the findings of the present experiments should be treated with caution. The reason we used the type of shearing pattern shown in Fig. 1 was to mimic the retinal motions present when a corrugation is translated and pursued. Rendering an exact orthographic projection of a rotating corrugation would not have suited our purposes. Had we done so, however, our findings might have been quite different. In this situation, depth order might be expected to remain ambiguous with or without knowledge of the direction of eye movement. However, we speculate that this may not be the case for orthographic projections of rotating objects such as spheres and cylinders. When these translate they resemble objects rolling across a surface. Even under orthographic projection, therefore, the direction of translation may bias the observer into seeing one particular depth ordering, congruent with the direction of roll. Corrugations, on the other hand, cannot roll and so are less likely to be interpreted in such a way.

4.3. Depth sensitivity

In Experiment 1 the slope of the psychometric function became steeper during eye movement, suggesting increased depth sensitivity in this condition. Why might this be the case? Slope is determined by the relationship between signal and noise. One explanation of the data is therefore that external noise (i.e. retinal jitter) decreased during eye movement. Work by Cornilleau-Peres and colleagues supports this view (see Cornilleau-Peres & Gielen, 1996 for review). For instance, Cornilleau-Peres and Droulez (1994) compared curvature discrimination

in conditions that differed in the quality of image stabilisation achieved by their observers. Performance was best in the object-rotation condition where stabilisation was best. Conversely, performance was worst in the object-translating condition because the pursuit eye movements were unable to stabilise the image as well. Intermediate findings were found for a third, head-moving condition. Though generally supportive of the external-noise hypothesis, it should be noted that their experiments differed from ours in two important respects. First, they did not require depth order to be disambiguated in their experiments (though we note in passing that some of their observers reportedly suffered depth reversals). Second, the sensitivity change we found was between stimuli whose retinal images were substantially different, whereas the differences they found were between stimuli that differed only in the degree of stabilisation achieved. Moreover, in an explicit test of the stabilisation hypothesis, van Damme and van de Grind (1996) compared curvature discrimination and motion detection in head-moving and head-stationary conditions. Intriguingly, while head-movements improved curvature discrimination in most of their observers, motion detection was made *worse*. This suggests the limiting factor may be internal and at a much later stage of analysis.

4.4. Relation to work on head movements

Examining how activity affects depth perception is certainly not a new idea. Crucially, however, much of the earlier work emphasises head translation (e.g. Ono & Steinbach, 1990; Rogers & Graham, 1979) and so differed from our work in many key respects. Gibson originally coined the term *motion perspective* to describe the parallax produced by a moving observer (Gibson, Olum, & Rosenblatt, 1955). The term is a good one because it emphasises the fact that head movements *generate* motion cues to depth. Eye movements, on the other hand, simply interfere with them. This may turn out to be an important difference when evaluating the type of heuristics used for the recovery of depth during these two different types of activity. For instance, recent work by Wexler and colleagues has shown that object stationarity is an important constraint on the number of possible interpretations of the motion created by a head movement. Moreover, stationarity may override the much-revered assumption of rigidity (Wexler, Lamouret, & Droulez, 2001; Wexler, Panerai, Lamouret, & Droulez, 2001). However, the same cannot be said for the situation studied here. The pursuit we examined was made to objects that moved independently of the observer, so the stationarity assumption does not apply.

Most authors assume the extra-retinal contribution during head movement is vestibular in origin (Rogers & Rogers, 1992). However, as both Freeman and Fowler

(2000) and Nawrot (2003) point out, lateral head movement is usually accompanied by a compensating eye movement, so both vestibular and eye-movement information could be used to obtain an estimate of head-translation velocity. Nawrot has gone one stage further and suggested that vestibular information is in fact ignored. He examined the perceived spatial phase of depth corrugations created by combining a shearing motion aftereffect and combinations of head and eye movements. Previous work by Ono and Ujike (1994) showed that this combination produces an impression of depth that depends not only on the ‘phase’ of the shearing aftereffect but also on the direction of head movement. By removing the need for optokinetic contribution to the compensating eye movement, Nawrot found observers were unable to report spatial phase unambiguously, despite the fact that disambiguating vestibular information was available to them. He achieved this by first estimating the gain of the vestibular–ocular reflex in the dark and then moving the test stimulus in such a way as to produce near perfect image stabilisation without the need for additional optokinetic compensation. To underscore his finding, Nawrot found that spatial phase was reported unambiguously in all other conditions that contained optokinetic components.

Even during head movement, therefore, extra-retinal, eye-velocity signals may be crucial for the judgement of depth from retinal motion.

Acknowledgments

The work was supported by the EPSRC. Some of the data has appeared in preliminary form at the Vision Sciences conference (2001) in Sarasota, FL, USA.

References

- Barnes, G. R. (1993). Visual-vestibular interaction in the control of head and eye movement: the role of visual feedback and predictive mechanisms. *Progress in Neurobiology*, *41*, 435–472.
- Bradshaw, M. F., & Rogers, B. J. (1999). Sensitivity to horizontal and vertical corrugations defined by binocular disparity. *Vision Research*, *39*, 3049–3056.
- 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.
- Carpenter, R. H. S. (1988). *Movements of the eyes*. London: Pion.
- Collewijn, H., & Tamminga, E. P. (1984). Human smooth and saccadic eye-movements during voluntary pursuit of different target motions on different backgrounds. *Journal of Physiology-London*, *351*(June), 217–250.
- Cornilleau-Peres, V., & Droulez, J. (1994). The visual perception of three-dimensional shape from self-motion and object motion. *Vision Research*, *34*(18), 2331–2336.
- Cornilleau-Peres, V., & Gielen, C. C. A. M. (1996). Interactions between self-motion and depth perception in the processing of optic flow. *Trends in Neurosciences*, *19*(5), 196–202.

- Domini, F., & Caudek, C. (1999). Perceiving surface slant from deformation of optic flow. *Journal of Experimental Psychology: Human Perception and Performance*, 25, 426–444.
- Domini, F., & Caudek, C. (2003). 3-D structure perceived from dynamic information: a new theory. *Trends in Cognitive Sciences*, 7(10), 444–449.
- Ebisawa, Y., Minamitani, H., Mori, Y., & Takase, M. (1988). New methods for removing saccades in analysis of smooth pursuit eye movement. *Biological Cybernetics*, 60, 11–119.
- Freeman, T. C. A. (1999). Path perception and Filehne illusion compared: model and data. *Vision Research*, 39(16), 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(7), 941–945.
- Freeman, T. C. A., Banks, M. S., & Crowell, J. A. (2000). Extraretinal and retinal amplitude and phase errors during Filehne illusion and path perception. *Perception and Psychophysics*, 62(5), 900–909.
- Freeman, T. C. A., & Fowler, T. A. (2000). Unequal retinal and extra-retinal motion signals produce different perceived slants of moving surfaces. *Vision Research*, 40(14), 1857–1868.
- Freeman, T. C. A., Harris, M. G., & Meese, T. S. (1996). On the relationship between deformation and perceived surface slant. *Vision Research*, 36(2), 317–322.
- 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.
- Gibson, J. J., Olum, P., & Rosenblatt, F. (1955). Parallax and perspective during aircraft landings. *American Journal of Psychology*, 68, 372–385.
- Harris, M. G. (1994). Optic and retinal flow. In A. T. Smith & R. J. Snowden (Eds.), *Visual detection of motion*. London: Academic Press.
- Harris, M., Freeman, T., & Hughes, J. (1992). Retinal speed gradients and the perception of surface slant. *Vision Research*, 32(3), 587–590.
- Hayashibe, K. (1991). Reversals of visual depth caused by motion parallax. *Perception*, 20(1), 17–28.
- Hogervorst, M. A., Bradshaw, M. F., & Eagle, R. A. (2000). Spatial frequency tuning for 3-D corrugations from motion parallax. *Vision Research*, 40, 2149–2158.
- Koenderink, J. J. (1986). Optic Flow. *Vision Research*, 26(1), 161–179.
- Lappe, M., Bremmer, F., & van den Berg, A. V. (1999). Perception of self-motion from visual flow. *Trends in Cognitive Sciences*, 3(9), 329–336.
- Liter, J. C., & Braunstein, M. L. (1998). The relationship of vertical and horizontal gradients in the perception of shape, rotation and rigidity. *Journal of Experimental Psychology: Human Perception and Performance*, 24(4), 1257–1272.
- Meese, T. S., & Harris, M. G. (1997). Computation of surface slant from optic flow: orthogonal components of speed gradient can be combined. *Vision Research*, 37, 2369–2379.
- Nawrot, M. (2003). Eye movements provide the extra-retinal signal required for the perception of depth from motion parallax. *Vision Research*, 43, 1553–1562.
- Ono, H., & Steinbach, M. J. (1990). Monocular stereopsis with and without head movement. *Perception and Psychophysics*, 48, 179–187.
- Ono, H., & Ujike, H. (1994). Apparent depth with motion aftereffect and head movement. *Perception*, 23, 1241–1248.
- Rogers, B., & Graham, M. (1979). Motion parallax as an independent cue for depth perception. *Perception*, 8, 123–134.
- Rogers, B. J., & Collett, T. S. (1989). The appearance of surfaces specified by motion parallax and binocular disparity. *Quarterly Journal of Experimental Psychology*, 41A(4), 697–717.
- Rogers, B. J., & Graham, M. (1982). Similarities between motion parallax and stereopsis in human depth perception. *Vision Research*, 22, 261–270.
- Rogers, S., & Rogers, B. J. (1992). Visual and nonvisual information disambiguate surfaces specified by motion parallax. *Perception and Psychophysics*, 52(4), 446–452.
- Royden, C. S., Banks, M. S., & Crowell, J. A. (1992). The perception of heading during eye-movements. *Nature*, 360(6404), 583–587.
- Turano, K. A., & Massof, R. W. (2001). Nonlinear contribution of eye velocity to motion perception. *Vision Research*, 41, 385–395.
- Ujike, H., & Ono, H. (2001). Depth thresholds of motion parallax as a function of head movement velocity. *Vision Research*, 41, 2835–2843.
- van Damme, W. J. M., & van de Grind, W. A. (1996). Non-visual information in structure-from-motion. *Vision Research*, 36(19), 3119–3127.
- Wallach, H., & O'Connell, D. N. (1953). The kinetic depth effect. *Journal of Experimental Psychology, Experimental-Psychology*.
- Wertheim, A. H. (1994). Motion perception during self-motion—the direct versus inferential controversy revisited. *Behavioral and Brain Sciences*, 17(2), 293–311.
- Wetherill, G. B., & Levitt, H. (1965). Sequential estimation of points on a psychometric function. *British Journal of Mathematical and Statistical Psychology*, 18(1), 1–10.
- Wexler, M., Lamouret, I., & Droulez, J. (2001). The stationarity hypothesis: an allocentric criterion in visual perception. *Vision Research*, 41, 3023–3037.
- Wexler, M., Panerai, F., Lamouret, I., & Droulez, J. (2001). Self-motion and the perception of stationary objects. *Nature*, 409, 85–88.