

Pre-print. Edited version is at Spatial Vision, In Press.

Detecting binocular 3-D motion in static 3-D noise: No effect of viewing distance.

Julie M. Harris* and Jane H. Sumnall[§]

**Department of Psychology, Ridley Building, University of Newcastle,
Newcastle upon Tyne NE1 7RU, U.K.**

Short title: 3-D motion and viewing distance.

***Corresponding author. E-mail: J.Harris@ncl.ac.uk Fax: +44 191 2225622**

[§]Current address: School of Psychology, Cardiff University, PO BOX 901, Cardiff CF10 3YG, U.K.

Abstract.

Relative binocular disparity cannot tell us the absolute 3-D shape of an object, nor its 3-D trajectory if it is moving, unless the visual system has independent access to how far away the object is at any moment. Indeed, as the viewing distance is changed, the same disparate retinal motions will correspond to very different real 3-D trajectories. In this paper we were interested in whether binocular 3-D motion detection is affected by viewing distance. We used a visual search task in which the observer is asked to detect a target dot, moving in 3-D, amidst 3-D stationary distractor dots. We found that distance does not affect detection performance. Motion-in-depth is consistently harder to detect than the equivalent lateral motion, for all viewing distances. For a constant retinal motion with both lateral and motion-in-depth components, detection performance is constant despite variations in viewing distance that produce large changes in the direction of the 3-D trajectory. We conclude that binocular 3-D motion detection relies on retinal, not absolute visual signals.

Introduction.

In principle, the visual system can obtain the direction of 3-D motion from the difference between the right and left eye retinal motion signals. Regan and colleagues have argued (for example see Regan, 1991; Beverley and Regan, 1973, 1975) that this is the only information required to judge whether a moving object will hit your head. However, the visual system needs an estimate of the viewing distance, (D), to obtain the real 3-D motion speed and direction, and to predict the actual 3-D trajectory in Euclidean space (see Cumming, 1994 for a review). This viewing distance estimate cannot be obtained from the horizontal relative disparity field alone (Howard and Rogers, 1995, suggest alternative ways to obtain the distance, including the possible use of vertical disparities). There is some indication that detection of depth from binocular disparity depends somewhat on viewing distance (Kedzia and Koczorowski, 1986; Foley, 1991; Bradshaw et al., 1995), but this is controversial (Ogle, 1958; Larson, 1987; Harris et al., 1998a).

We have previously studied suprathreshold binocular 3-D motion detection and compared lateral motion in the frontoparallel plane (x -motion) with motion directly towards or away from the observer (z -motion), for experimental situations in which the eyes receive equivalent magnitudes of retinal motion (Harris et al., 1998b; Sumnall and Harris, 1999a). We found that, in a visual

search task where a moving target dot is detected amidst a stationary field of random 3-D distractors, z-motion of the target is harder to detect than x-motion. Further, adding a little x-motion to z-motion enhances detection, whilst a considerably larger amount of z-motion must be added to x-motion to improve performance (Sumnall and Harris, 1999a).

Here, we systematically explore detection of x-, z- and x+z motion at 3 different viewing distances, to test whether the mechanisms involved in detection of binocular 3-D motion take account of the viewing distance. In other words, is it easier to detect the same retinal motion when it corresponds to a larger real z-motion (at a bigger viewing distance)? And if the real z-motion is large, do we need a correspondingly larger x-motion to aid detection? To summarise, we find that for a constant retinal motion, there is no change in performance as the viewing distance is increased from 75cm to 3m, despite the fact that this manipulation changes the real amount of z-motion 16-fold.

Methods.

Stimuli.

The stimuli were sparse random dot stereograms, generated and presented using a Pentium 200 MHz PC with fast 24 bits/pixel graphics, running at approximately 67 Hz. The pair of stereo half-images was displayed side-by-side on an Eizo 21 inch monochrome display monitor. A modified Wheatstone stereoscope was used to present each eye with a single half-image (after Johnston, Cumming and Parker, 1993). Observations took place in a darkened laboratory.

The stimuli were composed of a variable number of stationary bright dots (of luminance 96 cd/m^2), each of which had an independent disparity sample added (from a uniform distribution, ± 6 min arc around fixation), presented on a dark background. Thus, when the stereograms were fused, the dots appeared as randomly distributed throughout a notional end-on cylinder, centred on a small bright cross in the fixation plane. Stimuli were presented at one of three viewing distances (0.75m, 1.5m or 3m) and at each distance, the half-images subtended 2 degrees and each bright dot subtended 1 min arc.

On half of the presentations, the stimulus also contained a single moving target dot, identical to the bright stationary dots in every way, except that it moved throughout the presentation. The target dot moved along a chosen 3-D trajectory (undergoing a combination of motion in the x- and z- directions), but was always constrained to move within a small cylinder subtending 30 min arc, centred on the larger stimulus cylinder, and crossing the fixation plane midway through the presentation (see Figure 1 for a schematic representation of the stimulus).

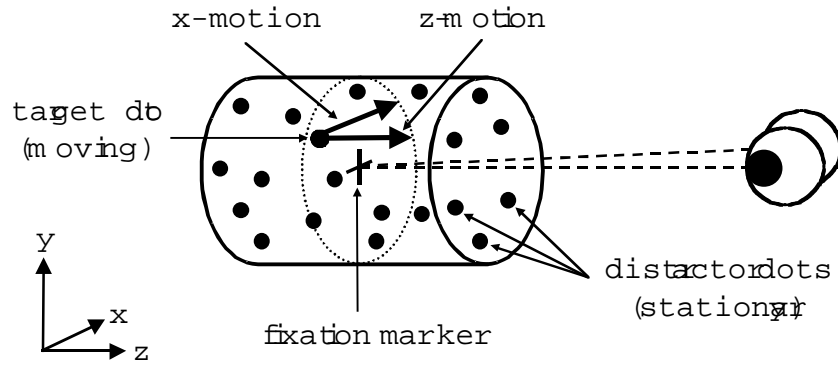


Figure 1: A schematic representation of the stimulus (Cyclopean view), showing a target dot defined by x-motion (motion in the frontoparallel plane) and z-motion (motion in the median plane with respect to the head).

Viewing distance manipulation and conditions.

When considering binocular 3-D motion, one can express motions in terms of the right and left eye retinal angular velocities (V_r and V_l) generated by a real 3-D motion, or by the real 3-D velocity itself (V). V can be decomposed into two orthogonal motion components, a z-motion component in a median plane with respect to the head, (V_z), and a x-motion component in a frontoparallel plane (V_x). Approximations for both V_z and V_x can be obtained in terms of the retinal motions, as follows:

$$V_z \approx (V_r - V_l)D^2/I \quad (1)$$

$$V_x \approx (V_r + V_l)D/2 \quad (2)$$

where D is the viewing distance and I is the inter-pupillary distance (see Sumnall and Harris, 1999a, for an alternative expression for Equation 1, and Regan, 1993, for a derivation of Equation 2).

The key point to take away from Equations (1) and (2) is that the x- and z-components of motion depend *differently* on viewing distance. V_z scales approximately with $1/D^2$, whilst V_x scales

approximately with $1/D$. Thus for constant retinal motions, at different viewing distances the actual 3-D motion represented will be quite different. Thus, constant retinal motion corresponds to different 3-D trajectories as viewing distance is changed.

For our experiment we chose to use a *constant* retinal motion, whilst manipulating viewing distance. This was achieved by moving observers further away from the screen, whilst scaling the physical stimulus to produce identical retinal stimuli. The stereoscope allowed variable vergence by adjustment of the central viewing mirrors, which were adjusted by each observer to provide comfortable fusion. We reasoned that subjects would adjust the mirrors so that the resulting vergence matched the cues to vergence provided by accommodation to the screen surface and by proximal vergence cues (e.g. knowledge about the size and distance of the monitor, see Judge, 1991).

We used three different distances (0.75m, 1.5m, 3m) for each of three stimulus conditions: target dots with only x-motion (left and right retinal motions are equal), target dots with only z-motion (left and right retinal motions are equal and opposite), and a target dot having a combination of the two orthogonal motions, x+z motion. Table I shows the real (absolute) trajectories generated by the same retinal motion at the three different viewing distances (calculated using Equations 1 and 2, with an inter-pupillary distance of 6cm).

Trajectory of motion	Retinal motions (min arc/s)		Real world (absolute) motions (cm/s)					
			D = 0.75m		D = 1.5m		D = 3m	
	V _r	V _l	V _x	V _z	V _x	V _z	V _x	V _z
(a) x-motion	2.251	2.251	0.049	0.000	0.098	0.000	0.196	0.00
(b) z-motion	2.251	-2.251	0.000	1.228	0.000	4.911	0.000	19.640
(c) x+z motion	4.503	0.000	0.049	1.228	0.098	4.911	0.196	19.640

Table I: The same retinal motion corresponds to different real world (absolute) motions as viewing distance is changed.

Table II shows how, for the combined x+z motion, changing the viewing distance results in a

very different angle of 3-D motion trajectory. The angular trajectories are measured from the median plane with respect to the head, such that z-motion has a constant angular trajectory of 0 degrees, whilst x-motion has a constant angular trajectory of ± 90 degrees. Thus the angular trajectory (α) of any motion can be calculated using:

$$\alpha = \tan^{-1}(V_x/V_z) \quad (3)$$

Trajectory of motion	Retinal motions (min arc/s)		Real world angular trajectory α (degrees)		
	V _r	V _l	D = 0.75m	D = 1.5 m	D = 3m
(c) x+z motion	4.503	0	2.285	1.143	0.572

Table II: The same retinal motion corresponds to different real world angular trajectories as viewing distance is altered.

Observers.

Two observers performed the experiments, one of the authors (JHS) and an experienced but naïve psychophysical observer, who was not aware of the experimental purpose. Both observers had normal or corrected-to-normal vision and were experienced in binocular visual experiments.

Psychophysical procedure.

A temporal two alternative forced choice procedure was used. On each trial the observer viewed two stimulus intervals (each lasting 1080ms and separated by an inter-stimulus interval of 540ms). In one (whose order was chosen at random from trial to trial), only the stationary distractor dots were present. In the other, the moving target dot was also present. The observer's task was to decide which interval contained the target dot. Each experimental run consisted of 100 trials. The percent correct for detection was recorded from at least 3 runs and a mean and standard error obtained.

Results.

As discussed in the Methods, the z- and x-components of motion scale differently with viewing distance. V_z is inversely proportional to the square of viewing distance, whilst V_x is inversely proportional to viewing distance. Thus, as viewing distance is altered, the same retinal motion can correspond to rather different real 3-D trajectories.

Here we tested whether viewing distance had a differential effect on the detection of z-motion and x-motion as the number of distractor dots was increased. Figure 2 shows percent correct for detection of the target dot as a function of the number of distractor dots for (a) x-motion, (b) z-motion and (c) x+z motion, for two observers. Each curve shows data for a different viewing distance.

Figure 2a shows percent correct detection for x-motion (different symbols indicate different viewing distances). There are no obvious differences between the curves for the three different viewing distances, suggesting little or no effect of viewing distance. Figure 2b shows percent correct detection for z-motion.

As we found previously (Harris et al., 1998b; Sumnall and Harris, 1999a), performance is poorer for z-motion than x-motion, despite there being the same magnitude of retinal motion in each eye, for the two conditions. Again, in Figure 2b the curves for the different viewing distances can be seen to closely overlap one another and there appears to be no effect of viewing distance.

Note again that since the absolute values of motion corresponding to the same retinal motions are very different for the different viewing distances, these data suggest that the visual system is responding to retinal, rather than absolute motion.

Next we considered what happens when x- and z-motions are combined. We found previously that adding a small amount of x-motion makes z-motion easier to detect (Sumnall and Harris, 1999a). But, for the same retinal motion, the z-motion component is much larger at 3m than at 0.75m. Is a larger amount of x-motion needed to produce the same improvement in performance? If so, then the same x-component of *retinal* motion should generate less of an improvement in

performance. Figure 2c shows the percent correct detection for combined x+z motion as a function of the number of distractor dots. Again, the viewing distance has very little effect: the curves for each observer closely overlap one another.

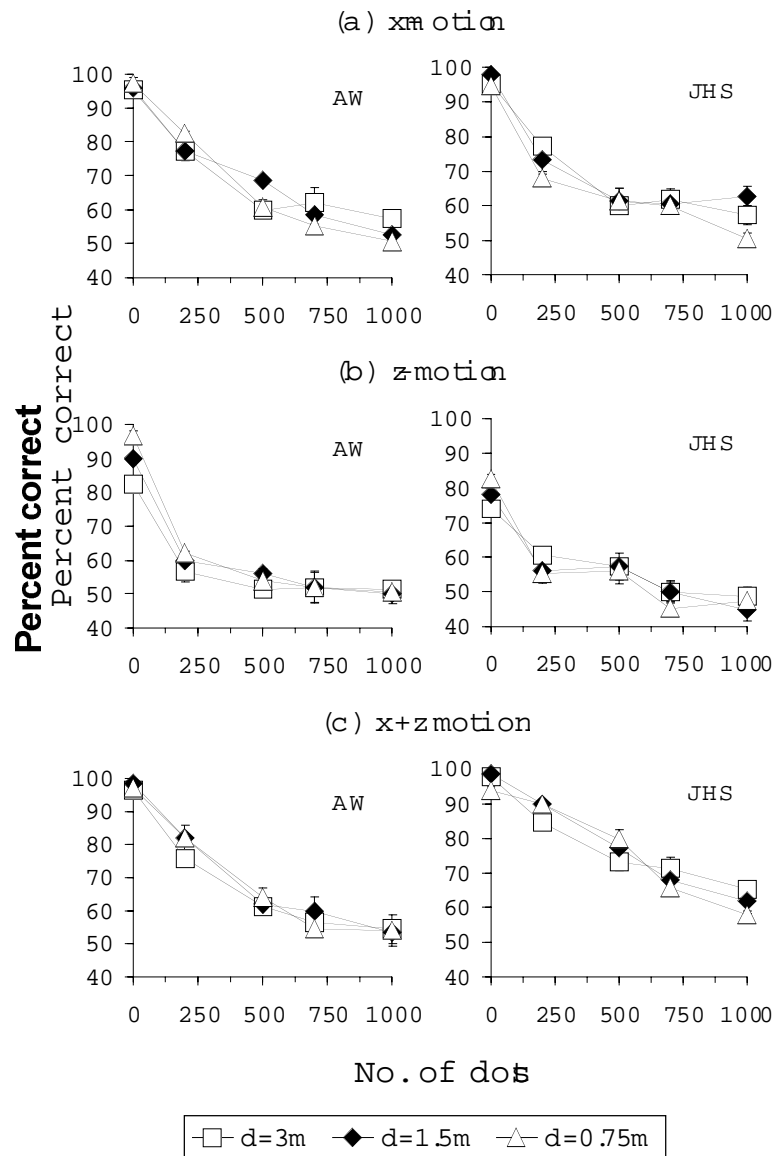


Figure 2: Percent correct detection of a moving target dot as a function of the number of distractor dots, for three viewing distances ($D=0.75\text{m}$, 1.5m , 3m). The target dot moves along different 3-D trajectories: (a) in the frontoparallel plane (x-motion); (b) in the median plane with respect to the head (z-motion); (c) along an intermediate trajectory (x+z motion). Each datum shows the mean percent correct from three experimental runs and error bars show the standard error of the mean.

Hence, for x-motion, z-motion and x+z motion, we find no evidence to suggest that viewing distance has any effect on performance. The data suggest that the visual system responds to a combination of retinal motion and makes no attempt to scale the motion for viewing distance.

Discussion.

Binocular 3-D motion detection does not change with viewing distance.

The results of the experiment reported here tested directly whether performance depended on viewing distance, and found that it did not. We have previously shown that z-motion is harder to detect than x-motion in a suprathreshold visual search task (Harris et al., 1998b; Sumnall and Harris, 1999a) and for 2-frame motion threshold (Sumnall and Harris, 1999b). One might expect a discrepancy between x-motion and z-motion if the visual system responds to real x- and z-motions, rather than retinal motions. What would we expect to find if this were the case? In our previous visual search experiments (Harris et al, 1998b; Sumnall and Harris, 1999a) we used a viewing distance of 3m, and values of x- and z-motion such that each eye viewed the same retinal speed (but of the same direction for x-motion, and the opposite direction for z-motion). The real z-motion is consequently much larger than the real x-motion (see Table I). If the visual system responded to an estimate of real distance moved, or real speed, then one would predict that the thresholds might be lower for z-motion than for x-motion. In fact the reverse is true. We suggested an explanation for this in terms of the existence of a specialised mechanism for the detection of z-motion that relies on detecting the rate of change of binocular disparity, but with low spatio-temporal resolution (see also Cumming and Parker, 1994). Thus our previous results cannot be explained by an explanation based upon viewing distance.

Viewing distance and binocular vision.

To obtain the veridical 3-D shape, or motion, or distance, of an object is not possible using only horizontal relative binocular disparity. In order to see the veridical shape, or obtain the veridical 3-D motion trajectory, at a variety of different viewing distances, the visual system must make some estimate of the viewing distance. In the real world, there are many possible cues to distance, including known size, vertical disparity, congruence between accommodation and vergence etc (see Foley, 1991 for a review of distance perception). However, in the laboratory, many studies suggest that such veridical information is not always available. For example, in 3-D shape perception, the same retinal stimulus can correspond to objects of very different shape

when viewed at different viewing distances. There is evidence from a 3-D shape perception study that, in a restricted situation where only binocular disparity, accommodation and vergence information was present, the visual system makes shape discrimination errors consistent with mis-estimating viewing distance (Johnston, 1991). Other studies, in which vertical disparity and other information is present (for example see Rogers and Bradshaw, 1993; Glennerster et al., 1998), suggest that in a richer environment, viewing distance can sometimes be obtained veridically. It is likely that in the real world there would be sufficient visual information available to obtain an estimate of object distance, and thus to convert motion estimates from retinal to real-world coordinates. Of course, under those circumstances, many other visual cues would be contributing to the perception of 3-D distance and motion.

Almost all studies of the sensitivity of the visual system to binocular disparity have assumed that the system is sensitive to relative retinal disparity and have not considered the functional consequences of this assumption. However, it is possible that for some tasks, the system might instead be sensitive to real depth, that has been scaled by an estimate of viewing distance. A viewing distance estimate could (at least in principle) be obtained from such cues as accommodation or convergence. Such ideas would predict that stereoacuity (measured in terms of relative disparity) would vary with viewing distance. Some studies have found this to be the case (Kedzia and Koczorowski, 1986; Bradshaw et al., 1995), but others have not (Ogle, 1958; Larson, 1987; Harris et al., 1998a), and it is not clear why these differing results have been found. Our study here suggests that, at least in the domain of 3-D motion detection, under conditions where only binocular visual information is available, the assumption that the visual system responds to *relative*, retinal information, is a sound one.

Acknowledgements.

This work was presented, in part, at the AVA Meeting on Visual Search, London 1999, and was supported by MRC (UK) grant G9533618N to JMH. We thank Ola Willis for her participation as an observer in these experiments.

References.

- Beverley, K. I. and Regan, D. (1973). Evidence for the existence of neural mechanisms selectively sensitive to the direction of movement in space. *Journal of Physiology* **235**, 17--29.
- Beverley, K. I. and Regan, D. (1975). The relation between discrimination and sensitivity in the perception of motion in depth. *Journal of Physiology* **249**, 387--398.
- Bradshaw, M. F., Rogers, B. J. and Glennerster, A. (1995). Are binocular disparity thresholds affected by viewing distance? *Investigative Ophthalmology and Visual Science* **36**, S230.
- Cumming, B. G. (1994). Motion-in-depth. In: *Visual Detection of Motion*. A. T. Smith and R. J. Snowden (Eds.). Academic Press, London, pp. 333--366.
- Foley, J. M. (1991). Binocular space perception. In: *Binocular Vision*. D. Regan (Ed.). CRC Press, Boca Raton, Fl., pp. 75--92.
- Glennerster, A., Rogers, B. J. and Bradshaw, M. F. (1998). Cues to viewing distance for stereoscopic depth constancy. *Perception* **27**, 1357--1365.
- Harris, J. M., Keeble, D. R. T., and Nishida, S. (1998a). Effects of viewing distance on stereoacuity, Panum's fusional limit and upper depth limit. *Investigative Ophthalmology and Visual Science* **39**, S623.
- Harris, J. M., McKee, S. P. and Watamaniuk, S. N. J. (1998b). Visual search for motion-in-depth: Stereomotion does not 'pop out' from disparity noise. *Nature Neuroscience* **1**, 165--168.
- Howard, I. P. and Rogers, B. J. (1995). *Binocular vision and stereopsis*. Oxford University Press, Oxford.
- Johnston, E. B. (1991). Systematic distortions of shape from stereopsis. *Vision Research* **31**, 813--826.
- Johnston, E. B., Cumming, B. G. & Parker, A. J. (1993). Integration of depth modules: stereopsis and texture. *Vision Research* **33**, 827--838.
- Judge, S. J. (1991). Vergence. In: *Vision and Visual Dysfunction: Eye Movements*, R. H. S. Carpenter, (Ed.). MacMillan, London.
- Kedzia, B. and Koczorowski, P. (1986). Stereopsis angle in relation to proximity and induced refractive errors. *American Journal of Optometry and Physiological Optics* **63**, 432--441.

Larson, W. (1987). Letter to the editor: the effect of test proximity and induced refractive error on stereoacuity. *American Journal of Optometry and Physiological Optics* **64**, 377--378.

Ogle, K. N. (1958). Presentation of our knowledge of stereoscopic vision. *A. M. A. Archives of Ophthalmology* **60**, 755--774.

Regan, D. (1991). Depth from motion and motion-in-depth. In: *Binocular Vision*. D. Regan, (Ed.). CRC Press, Boca Raton, Fl., pp. 137--169.

Regan, D. (1993). Binocular correlates of the direction of motion in depth. *Vision Research* **33**, 2359--2360.

Rogers, B. J. and Bradshaw, M. F. (1993). Vertical disparities, differential perspective and binocular stereopsis. *Nature* **361**, 253--255.

Sumnall, J. H. and Harris, J. M. (1999a). Binocular 3-D motion: Contributions of lateral and stereomotion. Submitted to *JOSA A*.

Sumnall, J. H. and Harris, J. M. (1999b). Minimum displacement thresholds for 3-D motion depend on the average binocular signal. *Investigative Ophthalmology and Visual Science* **40**, S765.