Speed Discrimination of Motion-in-depth Using Binocular Cues

JULIE M. HARRIS,*† SCOTT N. J. WATAMANIUK*

Although it is well known that motion-in-depth can be detected using binocular cues, it is not known whether those cues can be used to judge the speed of an object moving in depth. There are at least two possible binocular cues that could be used by the visual system to calculate three dimensional (3-D) speed: the rate of change of binocular disparity, or a comparison of the speeds of motion in the two eyes. We tested which of these cues is used to discriminate the speed of motion-in-depth. First, speed discrimination was measured for a dot moving away from the observer in depth (along the z-axis) and for a random dot stereogram in which a central square moved away from the observer in depth. These stimuli contained both disparity and monocular motion cues. Speed discrimination thresholds were as good for 3-D motion as for monocular sideways motion. Second, a dynamic random dot stereogram (in which the random dot pattern was replaced by a new dot pattern every frame) was used to remove consistent monocular cues. 3-D speed discrimination was now very poor, suggesting that the rate of change of disparity is not a good cue for 3-D speed. Finally, we tested whether observers were able to use the monocular motion cue from one eye to perform the speed discrimination task, or whether there had to be a comparison of the two eyes' monocular cues. By adding a small x-axis velocity component (with random direction) to the z-axis motion, it was possible to disrupt the monocular motion signals without altering the speed of the motion in 3-D. This manipulation did not disrupt the observers' performance, suggesting that monocular speed cues were not being used independently but that there was a comparison of monocular motion signals from the two eyes.

Motion perception   Speed discrimination   Stereopsis   Binocular vision   Depth perception

INTRODUCTION

The study of three dimensional (3-D) motion can be split into two general areas, those involving monocular cues (such as rate of expansion, see Regan & Beverley, 1978a, b) and those involving binocular cues (such as stereopsis, Julesz, 1971), which are the subject of this paper. Previous studies of motion-in-depth using binocular cues have been concerned with the conditions need for the *detection* of motion-in-depth (Julesz, 1971; Tyler, 1971; Regan & Beverley, 1973a; Cumming & Parker, 1994), or with direction discrimination of motion-in-depth (Beverley & Regan, 1973). This paper will consider what binocular cues are needed to judge the *speed* of motion directly towards or away from the observer.

Although the world is 3-D, the three dimensions are not equivalent at the input stages of the visual system. For the x and y-directions (the frontoparallel plane), the positions and motions of objects in the world are projected directly onto the retina. In the projection from

world to retina, information about position-in-depth and motion-in-depth (the z-direction) is lost. To retrieve z-position information, the visual system is able to compare the slightly different views from the two retinae and extract information about depth from the differences between them (Wheatstone, 1838): this is binocular stereopsis.

An important question is how the visual system extracts information about *motion* in the z-direction. At least two binocular cues could be used for the extraction of the speed of z-direction motion (Cumming & Parker, 1994; Regan, 1993). First, the visual system might use the rate of change of retinal disparity (the 3-D position cue) to judge the direction and speed of motion-in-depth. The disparity (δ) of a point P is defined as the difference between the angles subtended between P and the point of fixation (F) for each eye. From the schematic in Fig. 1:

$$\delta = \theta_R - \theta_L. \qquad (1)$$

Simple geometry determines that $\theta_R - \theta_L = \alpha - \beta$. If the angles α and β are small ($D, Z \gg I$) and the point P is not far from the mid-line, they are given approximately by: $\alpha = I/D$ and $\beta = I/Z$, where I is the interocular separation, D is the viewing distance and Z is the

---

*The Smith-Kettlewell Eye Research Institute, 2232 Webster Street, San Francisco, CA 94115, U.S.A.
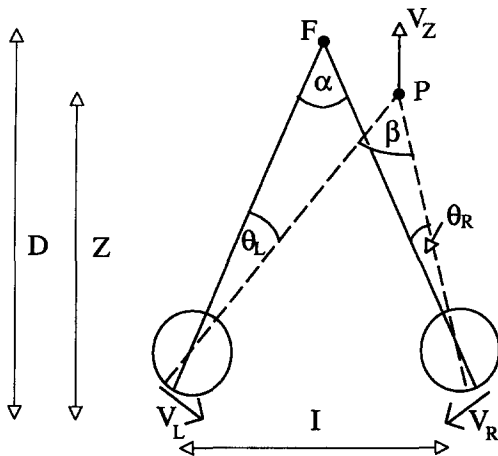†To whom all correspondence should be addressed.

FIGURE 1. An observer fixates on a point F. An object at point P moves away from the observer with speed V_z, causing the point's retinal projection to move with approximately equal and opposite speeds on each retina (V_L and V_R). The disparity of point P with respect to the fixation point, F, will also vary over time.

distance along the z-axis to point P. By substituting for α and β in equation (1), the following well-known approximate expression is obtained for disparity:

$$\delta \approx I(D - Z)/D^2. \tag{2}$$

To find the speed of motion in the z-direction, equation (2) can be differentiated and rearranged to give:

$$V_z \approx \frac{d\delta}{dt} D^2/I. \tag{3}$$

A second potential cue to the speed of motion-in-depth is a comparison of the speeds of motion in the two retinae. By differentiating equation (1) with respect to time:

$$\frac{d\delta}{dt} = \frac{d\theta_R}{dt} - \frac{d\theta_L}{dt} = V_R - V_L, \tag{4}$$

(where V_R and V_L are motions on the right and left retina, respectively) and substituting into equation (3):

$$V_z \approx (V_R - V_L)D^2/I. \tag{5}$$

Equations (3) and (5) are mathematically equivalent but they suggest potentially different ways in which the visual system might recover the speed of motion-in-depth.

Stimuli containing both these cues have been shown to generate a percept of motion-in-depth. Tyler (1971) measured the minimum displacement in depth needed to detect motion-in-depth. In a series of papers, Regan and Beverley studied binocular motion-in-depth for simple stereo stimuli (Regan & Beverley, 1973a, b; Beverley & Regan, 1973, 1974, 1975). They were the first to suggest that a comparison (they specifically suggested a ratio) of the motion signals from the two eyes could be used to find the direction of motion-in-depth. However, their stimuli contained both rate of change of disparity and monocular motion cues and thus could not distinguish between the two possible cues.

Julesz (1971) designed a stimulus in which there was rate of change disparity but no consistent motion cues and demonstrated that disparity cues alone were sufficient for the perception of motion-in-depth (although he did not explicitly measure thresholds for detection). In his stimuli, each frame of the dynamic random dot stereogram was defined by a new pattern of random dots, while from frame-to-frame the disparity of a central square increased. Thus, there was a changing disparity but no monocular motion cues.

A more recent study (Cumming & Parker, 1994) specifically addressed whether a rate of change of disparity was necessary, as well as sufficient, for the detection of motion-in-depth. Cumming and Parker showed that stimuli containing only disparity cues gave detection thresholds as low as stimuli containing both disparity and motion cues. By using the stimuli beyond the spatial and temporal range of stereopsis, they also showed that stimuli without consistent disparity cues did not generate a percept of motion-in-depth.

So far then, the evidence is in favour of the rate of change of disparity being an important cue to motion-in-depth. However, in all the experiments mentioned above, observers were either asked to detect motion-in-depth or to discriminate its direction (towards or away). Both of these tasks could be solved by using pure disparity cues and would not necessarily depend on a specialised system responding to motion. For instance, if asked in which direction the motion-in-depth was seen, the observer would only require information from the first and last frames (static disparity cues) and would need to know the temporal order of the static signals. In order to explore whether the visual system contains binocular mechanisms responsive to 3-D motion, it is more appropriate to use a task that reflects motion processing, such as speed discrimination.

The aim of this paper was to find out whether there is a system responsive to the *speed* of motion-in-depth, as has been established for the speed of 2-D motion (McKee 1981; Orban, de Wolf & Maes, 1984), and whether it requires the use of the rate of change of disparity or a comparison of monocular speeds, or both. We ask three main questions here:

(1) Is speed discrimination as good for motion-in-depth as for the equivalent monocular motion (motion seen when one eye is closed)?

(2) Is rate of change of disparity the basis for 3-D speed discrimination?

(3) Is a comparison of monocular motion signals from the two eyes the basis for 3-D speed discrimination?

In summary, we found that: (1) given both disparity and monocular motion cues, speed discrimination is as good for motion-in-depth as it is for the equivalent monocular motion. (2) When the stimulus contains only binocular disparity cues, speed discrimination thresholds are very high, suggesting that rate of change of disparity is a poor cue for 3-D speed discrimination. (3) A comparison of monocular motion cues is used to discriminate the speed-in-depth, rather than a

single monocular speed cue. This suggests that the visual system contains a binocular system responsive to the speed of motion-in-depth. These results have been presented in preliminary form elsewhere (Harris & Watamaniuk, 1994a).

EXPERIMENT I

For two dimensional (2-D) monocular motion, speed discrimination can be very good under some conditions (McKee, 1981; Orban *et al.*, 1984). Here, we tested whether speed discrimination for motion in the z-direction (directly towards or away from the observer) is as good as that for monocular motion. Motion of a dot in the z-direction requires that the dot moves in opposite x-directions on the two retinae (see Fig. 1). The appropriate comparison between the 3-D and 2-D conditions is to compare the binocular motion-in-depth with the monocular motion seen by one eye during the z-motion. Two experimental conditions were tested here. In the first, speed discrimination was measured for binocular motion-in-depth. In the second experimental condition we measured monocular speed discrimination by showing the observer one eye's half-image from the stereo stimulus.

Methods

Apparatus and stimuli. An Amiga 3000 was used to generate stereo stimuli which were presented on two $x - y$ CRT screens with P4 phosphor. Observers viewed the screens from 1.5 m via polarizors and a beam-splitter so that each eye saw only one screen. Stimuli were composed of light dots presented on a dark background and were viewed in a dimly lit room. By plotting a matrix of non-overlapping dots (with center-to-center spacing of 1.6 min arc), it was possible to measure their space-averaged luminance. Measured through the polarizors and beam-spitter, the dot luminance was $6 \, cd/m^2$. The luminance of the background was very low and could not be measured using the available apparatus (a hand-held Minolta luminance meter).

There were two kinds of stimuli used in this experiment. The first was a single dot moving in depth on a dark background and positioned at an angle of 16 min below a stationary fixation dot (see Fig. 2a). The initial horizontal position of the dot was chosen at random to be between ± 12 min arc on either side of the fixation position. The second stimulus was a sparse random dot stereogram, made up of 200 dots and subtending an angle of 1.5 deg containing a central square of width

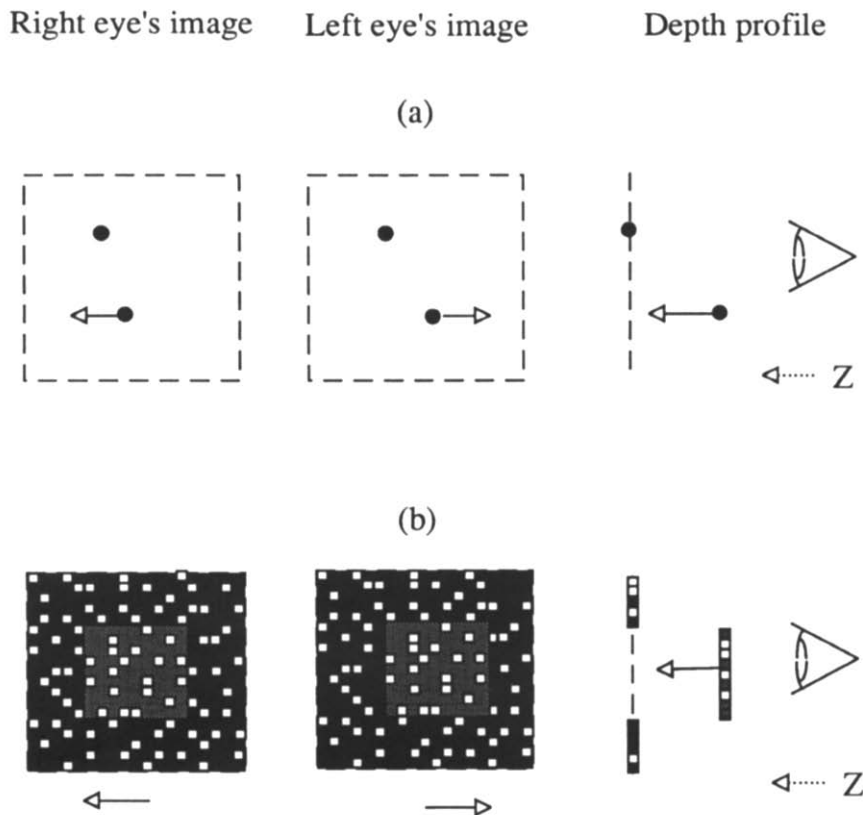Right eye's image   Left eye's image   Depth profile

(a)

(b)

FIGURE 2. (a) Schematic illustration of the single moving dot stimulus. The first panel shows the right eye's screen, on which there was a stationary white fixation dot on a dark background (shown as dark on light here for convenience) and a dot moving to the left below the fixation dot. The next panel shows the left eye's screen, on which the lower dot moved to the right. When presented stereoscopically, the lower dot moved away from the observer, as shown in the third panel. (b) Schematic illustration of the random dot stereogram stimulus. The stereogram was made up of white dots on a dark background. The dots in a central square region moved in depth (here, the central square area is shown as dark grey and its motion is shown by the arrows below the stimulus). On the right eye's screen a central square area moved to the left. On the left eye's screen it moved to the right. When presented stereoscopically, the square moved away from the observer in depth.

0.75 deg, which moved in depth with respect to the stationary surround (see Fig. 2b). Initially the dot (or central square in the random dot stereogram) was given a crossed disparity to simulate a depth of n cm (typically 14 cm, a disparity of 12.8 min arc) in front of the reference (i.e. closer to the observer than the reference). This "static portion" of the stimulus was displayed for 480 msec. During the 900 msec (on average) of the "moving portion" of the stimulus, the dot moved continuously away from the observer until it had an uncrossed disparity placing in n cm behind the reference, at which point the stimulus disappeared and was replaced by a blank screen.

Motion-in-depth was simulated by varying the horizontal positions of elements (and thus the stereo disparity) on the screens from frame to frame. For instance, to simulate a dot moving away from the observer, the dot on the right eye's screen should move to the right and the dot on the left eye's screen should move to the left. The horizontal positions of the elements on the right and left eye's screens were calculated using the following equations:

$$x_1 = \frac{(Z - D)I}{2D} + X; \quad x_r = \frac{(D - Z)I}{2D} + X. \quad (6)$$

Such positioning resulted in an equation for disparity given by:

$$\delta = \frac{x_r - x_1}{D} = \frac{I(D - Z)}{D^2}. \quad (7)$$

These equations are similar to those used for ray-tracing an image to give an exact simulation of the left and right eyes views of a real object. However, they do not exactly correspond to those equations because we did not want to include perspective cues (e.g. as a square moves away from the observer its elements appear to move closer together). For the random dot stereogram stimulus, it was ensured that the background dots formed a "window" through which the target dots moved. As dots were occluded by the background, new dots were randomly positioned to replace them at the opposite edge of the window. Images were presented at a frame rate of 50 Hz.

In the monocular condition, the same stimuli were used as for the motion-in-depth conditions, except that observers were allowed to view only the screen presenting the right eye's half-image (thus they saw sideways motion, rather than motion-in-depth). The same arrangement of apparatus was used, with the images presented via a beam-splitter and polarizors, the only difference being that the left eye saw a dark blank screen, rather than a stereo half-image.

Procedure. The observers were shown two intervals. In each they saw the static portion of the stimulus, followed by the moving portion. They were asked to decide in which interval the dot moved faster in depth. Each trial contained one interval in which a stimulus with a "standard" speed was shown (typically 0.32 m/sec in depth) and one in which a "test" stimulus, randomly chosen to have one of five speeds spaced evenly around

the standard, was shown. The order of presentation of the test and standard was also chosen randomly from trial to trial. In a single run of the experiment there were 100 trials. The percentage of trials where the observer saw the test as faster than the comparison was recorded and fitted with a cumulative normal using Probit analysis (Finney, 1971). A Weber fraction ($\Delta v/v$) was calculated where $\Delta v + v$ was the speed corresponding to the 75% point of the fitted function and v was the mean speed (corresponding to the 50% point).

Observers wee instructed to fixate a stationary reference dot throughout the stimulus presentation, and were specifically told not to track the moving dot (or dots). However, observers reported that it was very difficult not to track the motion-in-depth. Because eye movements were not monitored, we do not know the extent to which the eyes moved. However, even if there was considerable tracking, there would be relative motion between the tracked dot and the previously stationary "fixation" dot, and thus there would still be useful motion signals on the retinae.

It was important to be sure that observers were responding to the speed of motion and not to other cues (see McKee, 1981). For example, observers might respond to the size of the static disparity at the beginning or end of a trial (or the total distance moved in the monocular condition). To make such cues less useful to the observer, we randomised the duration of each "test" stimulus so that the larger disparities no longer necessarily corresponded to the faster speeds. It must be emphasised that such a manipulation does not prevent observers from using position or duration cues. Instead, it ensures that if such cues are used, performance will be poor. We used five durations, spaced in 130 msec steps around the mean duration of 900 msec. If observers were using the disparity cue, we would expect them to choose the longer duration (and hence larger disparity) stimuli as faster more often than the shorter duration (smaller disparity) stimuli.

We were restricted to a very narrow range of speeds in this experiment. If the speed were too slow, monocular speed discrimination would be expected to be poor (McKee, 1981; Orban *et al.*, 1984), but too fast a speed resulted in the stimulus being seen as diplopic for part of its motion. We wanted to maintain binocular fusion because we were specifically interested in whether the rate of change of binocular disparity was a cue to speed-in-depth. The speed of motion used in these experiments ranged from 0.26 to 0.38 m/sec in the z-direction (preliminary studies showed that speed discrimination was very poor for both monocular and binocular conditions when the speed was below 0.1 m/sec). This resulted in retinal speeds of about 0.2–0.3 deg/sec [this can be calculated from equation (5)]. Even at these speeds, observers reported diplopia at the beginning and end of some trials (those of a long duration, fast speed and hence large starting disparity).

Observers. The observers were the two authors and one other, who was naive as to the purpose of the

experiments. All observers had normal or corrected to normal vision and were highly experienced at performing psychophysical tasks. Two of the observers (JMH and HSS) were not previously experienced at speed discrimination tasks and were trained until their performance reached an asymptotic level.

Results and discussion

Figure 3 shows the Weber fractions found for speed discrimination of motion-in-depth (the binocular stimulus—dark bars) and of the 2-D motion seen when the observer saw one eye's image (monocular stimulus—light bars). Results are shown for both the random dot stereogram (in which a central square moved away from the observer in depth) and for a single dot, moving away from the observer in the presence of a stationary fixation dot. Very similar Weber fractions were found for both the random dot stereogram and single dot stimuli under both binocular and monocular conditions (although one observer, HSS, showed a slight tendency to perform better for the monocular task in one condition). Thus, observers are able to perform accurate discrimination of speed-in-depth, giving Weber
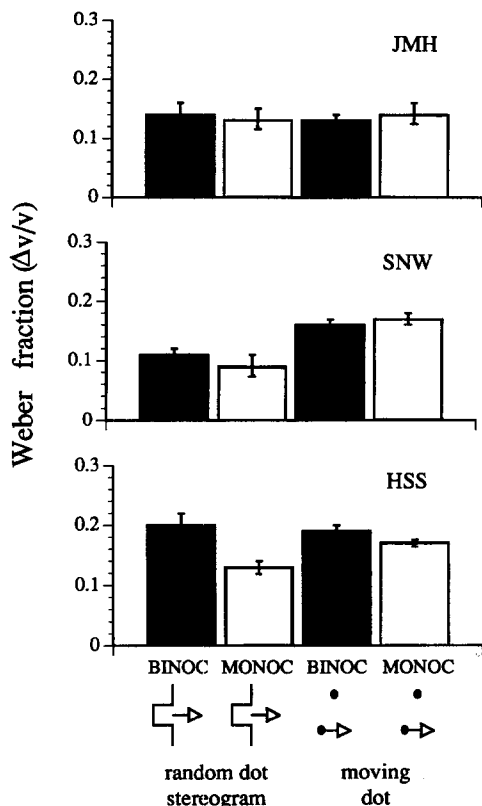

FIGURE 3. Weber fractions ($\Delta v/v$) for speed discrimination in four experimental conditions: (1) solid bars: motion-in-depth defined by a random dot stereogram, (2) open bars: monocular (sideways) motion of random dots when the observer was shown the right eye's stereo half-image, (3) solid bars: motion-in-depth of a single dot, (4) open bars: monocular (sideways) motion of a single dot when the observer was shown the right eye's stereo half-image. Error bars show SEs calculated from at least three experimental runs of 100 trails. The Weber fractions were very similar for all conditions. The speed of motion-in-depth was discriminated as easily as the speed of monocular motion.

fractions between 0.1 and 0.2 (in other words, they need an increase in speed of between 10% and 20% to see the test speed as faster on 75% of occasions). In addition, performance was as good for the motion-in-depth stimuli as for the equivalent monocular stimulus. This suggests that the mechanism for finding the speed-in-depth may be closely related to that for finding monocular speed. Weber fractions of 0.1–0.2 might appear rather high compared with those measured previously (Weber fractions of 0.05) for 2-D motion (McKee, 1981). Note however, that the monocular motion on each eye's retina was around 0.25 deg/sec, a speed too low to expect optimal monocular speed discrimination performance (as shown by McKee, 1981; Orban *et al.*, 1984).

In a control condition observers saw the initial disparity but no motion. They were asked in which interval the stimulus was at a greater depth. Because of the randomisation of the disparity (due to the duration randomisation for the original experiment), the observers' overall performance followed a pattern related to the duration that had been used in a particular trial. If these static cues were used to perform the speed discrimination task, we would expect similarly poor performance and a similar pattern of results. For the speed experiments, no trends with duration were seen, indicating that observers were not using static disparity cues.

A second control experiment tested whether observers could use static cues from the first and last frames, combined with an estimate of the interval between them, to "calculate" the speed-in-depth. Observers were shown the first frame for 200 msec, then a blank screen for the appropriated duration for a particular speed, and then the last frame for 200 msec. The observers were very poor at this task, performing at chance over the whole range of test speeds.

It may seem surprising that thresholds are as low for the motion-in-depth speed discrimination as for monocular speed discrimination (see Fig. 3). One might expect the motion-in-depth thresholds to be worse because they are based on the combination of two (presumably noisy) motion signals, whereas in the monocular condition the thresholds are based on only one motion signal. However, if the limiting noise was beyond the stage at which the monocular signals were combined, or if the noise from each monocular signal was not independent, similar thresholds might be expected.

The results shown in Fig. 3 are very different from those relating to the *detection* of motion-in-depth (Tyler, 1971). Tyler measured the threshold for the detection of motion-in-depth and found that when observers could no longer see the binocular motion, they still reported seeing motion if one eye was closed (the monocular condition). Such "stereomotion suppression" (Tyler & Foley, 1974) was not observed here. A possible reason might be that the suppression occurs close to detection threshold. The present experiment was performed using motions that were well above the detection threshold.

EXPERIMENT II

In Experiment I, it was established that observers can discriminate the speed of motion-in-depth as well as they can discriminate the speed of monocular motion. In the next two experiments we set out to find what cues are used to perform the discrimination.

When a random dot stereogram is used to generate static depth over a region of the stimulus, there are no cues to the position of the region in one stereo half-image (Julesz, 1971). However, for a region moving in depth, when one eye is closed the region is seen undergoing horizontal motion. In other words, a random dot stereogram depicting a square moving in depth contains both disparity cues and monocular motion cues. Either of these cues could have been used to discriminate the speed of the motion-in-depth in Experiment I.

To isolate the rate-of-change-of-disparity cue, Julesz (1971) designed the *dynamic* random dot stereogram, in which the disparity of the central square varied systematically from frame-to-frame, while on each frame the depth was portrayed by a new pattern of random dots. When an observer is allowed to view one stereo-half, only random motion is seen.

In this experiment we used a *dynamic* random dot stereogram to test whether the rate of change of disparity is a good cue to the speed of motion-in-depth.

Methods

The stimuli were similar to the random dot stimuli used in Experiment I except that here, on every new frame, a new pattern of random dots was generated to define the stereo stimulus. This allowed the presentation of a stimulus in which the depth of the central square changed but in which there were no consistent monocular motion cues. There were two experimental conditions, one in which observers saw the central square move in depth and a monocular condition where the observers saw only the right eye's image. For the mon-

ocular condition, it was expected that since there would not be any consistent monocular motion, the observers would be unable to perform the task.

As in Experiment I, observers were shown two intervals and asked in which interval the central square moved away faster. As before, there were five "test" speeds equally spaced around a "standard" speed, and five randomly chosen stimulus durations. The percentage of trials in which the observer reported the test stimulus to be the faster stimulus was recorded and the data were fitted with a cumulative normal using Probit analysis. Weber fractions were compared with those from Experiment I.

Results and discussion

Figure 4 shows the Weber fractions for the motion-in-depth condition, for the stimulus containing only the disparity cues (patterned bars) compared with those for the random dot stereogram stimuli from Experiment I, where both disparity and monocular motion cues were present (dark bars). When there were only disparity cues and no monocular motion cues, Weber fractions were very large. This suggests that the rate of change of disparity is not useful for judging the speed of 3-D motion. When performance was examined as a function of stimulus duration, rather than speed, it was found that observers tended to choose the longer duration stimuli as "faster" and the shorter duration stimuli as "slower". This suggests that observers were using a position or static disparity cue as well as, or instead of, a speed cue. It is therefore possible that the speed cue was not merely poor, but that it may not have been used at all by the observers.

For the monocular condition, observers saw only inconsistent random motion. We have not shown the data in Fig. 4 because observers were unable to perform the speed discrimination task, performing at around 50% correct (chance) for all combinations of test and standard. This confirms that there were no consistent monocular motion cues in the stimulus.
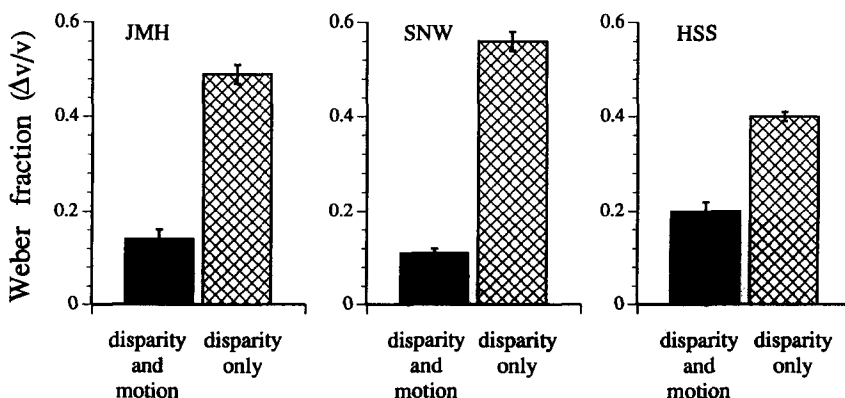

FIGURE 4. Weber fractions ($\Delta v/v$) for speed discrimination of motion-in-depth for two conditions: (1) dark bars: random dot stereogram stimulus containing both disparity and monocular motion cues, (2) patterned bars: dynamic random dot stereogram stimulus containing disparity cues but no consistent monocular cues. Error bars show SEs calculated from at least three experimental runs of 100 trails. The Weber fractions were very high for the stimulus without consistent monocular motion cues, suggesting that the rate of change of disparity (the only useful cue in the stimulus) is a very poor cue to the discrimination of speed-in-depth.

The results of this experiment may appear surprising because, as Julesz showed, a stimulus containing only changing disparity gives a clear percept of a square moving in depth. Furthermore, our observers had no difficulty in identifying the direction of object motion. However, the judgement of direction does not necessarily require *motion* to be processed. The observer has only to note the depth of the square at several times during its motion to see that it has moved in depth, a task requiring only the ability to detect the relative depth, which can be done using static disparity cues.

EXPERIMENT III

In the previous two experiments, we showed that observers can discriminate the speed of motion-in-depth, but not by using stereo cues alone. This suggests that a combination of monocular speed cues may be used to discriminate speed in the z-direction. However, it is possible that observers have independent access to one (or both) of the monocular motion cues and do not have to combine them to perform the speed discrimination task. Here, an experiment was performed to test whether the use of a single monocular cue (whether from a particular eye or chosen randomly from the two eyes) would give different results from the use of a *combination* of both monocular speed cues.

Methods

The stimuli used were similar to the single-moving-spot stimuli of Experiment I (see Fig. 5a). Here, in each interval, a very small x-axis motion was either added (motion to the right) or subtracted (motion to the left). In either case, the *3-D speeds* had the same evenly spaced

distribution around the standard speed as was used in Experiment I. Now, however, the *monocular* speed distributions would depend on whether the x-axis motion was added or subtracted. By adding x-axis motion to the test and subtracting x-axis motion from the standard, we were able to skew the monocular speed distributions so that the monocular speeds were either all below or all above the standard speed (see Fig. 5).

Figure 5b shows the condition in which the additional x-axis motion was added to *both* the test and standard stimulus, making the right eye motion vector 30% faster and the left eye motion vector 30% slower. We call this the "same" condition. Whether the observer used a single monocular cue (considering only the left, or only the right eye's vectors) or a combination of the motion from the two eyes, the same pattern of results would be expected as that in Fig. 5a because both monocular and 3-D speeds were spaced evenly around the standard stimulus. The same responses would also be expected for the case where the additional x-axis motion was subtracted from both test and standard (not shown).

Figure 5c shows the case where the additional x-axis motion was added to the test but subtracted from the standard. We call this the "different" condition. Now, if the observer used only the left eye, she would always see the test as slower because all the test speeds in the left eye are slower than the left eye's standard speed. Using the right eye, she would always see the test as faster because, for the right eye, the test speeds were always faster than the standard. If the eye were chosen randomly each trial, the observer would respond at the 50% level (chance) on average, for all test speeds. However, if the two monocular motion cues were combined to estimate 3-D speed, the observers response would be expected to be similar to the responses in

STIMULUS				RESPONSE		
standard		test		use	use	use combination
left eye	right eye	left eye	right eye	left eye	right eye	of two eyes
(a)				faster same slower	faster same slower	faster same slower
(b)				faster same slower	faster same slower	faster same slower
(c)				slower slower slower	faster faster faster	faster same slower

FIGURE 5. Diagram illustrating the stimulus speed distributions and expected responses for a single dot moving in the z-direction (a) or when the motion also contained a component in the x-direction (b and c). (a) Pure z-direction motion. The test stimuli were distributed evenly around the standard stimulus for monocular and 3-D motion. Thus the same pattern of responses would be predicted whether observers used only the left eye, only the right eye or a binocular combination giving 3-D motion. (b) x-direction motion added to both the test and standard stimuli. The test stimuli were again distributed evenly around the standard stimulus for monocular and 3-D motion. Again, the responses would be the same whether monocular cues or binocular comparison was used. (c) x-direction motion added to the test but subtracted from the standard. For the left eye, the test stimuli were always slower than the standard. For the right eye, the test was always faster than the standard. The 3-D test speeds were evenly distributed around the standard. If observers used monocular cues to perform the speed discrimination, a different pattern of results would be expected than if they were using a binocular combination of speed cues that gave 3-D speed.

Fig. 5a, because the 3-D test speeds were spaced evenly around the 3-D standard speeds. Thus, observers should respond similarly to the stimuli illustrated in Figs 5a and 5c if they used a combination of monocular speed signals.

In this experiment there were four possible arrangements of test and comparison, the two shown in Figs 5b, c and their mirror-images. A single experimental run contained equal numbers of the four stimulus arrangements presented in pseudo-random order. Data were stored separately for the four conditions. We also collected data for a monocular condition, in which the observers used only their right eye.

Results and discussion

The percent of "test faster" responses are plotted as a function of the z-direction speed of the test stimulus in Fig. 6a. In Fig. 6b we plot the percent of "test faster" responses for the monocular condition, as a function of the monocular speed that is equivalent to the z-axis motion in Fig. 6a (note that the *monocular* speeds were the same in the motion-in-depth and monocular conditions, the only difference being that observers saw only the right eye's stereo half-image in the monocular condition). The arrows show the speed of the standard stimulus. The percent correct is given in units of d' so that the psychometric functions can be fitted with straight lines. For each observer there is a graph showing the results for the binocular (motion-in-

depth) condition (upper graphs) and the monocular controls (sideways motion—lower graphs). In each graph we plotted the average of the two "same" conditions and the average of the two "different" conditions. The results from the "same" and "different" conditions are shown by the filled squares and open circles, respectively. The best fitting lines through the data are shown.

First, consider the plots for the monocular control condition (lower graphs), where the observers used only their right eye. When both the test and standard stimulus had the same additional x-axis motion added, observers responses took the form of a typical psychometric function with d' ranging from about -2 to $+2$ (roughly 10–90% correct). For instance, when the test speed was faster, observers would respond "test faster" on close to 100% of occasions, when slower, they would respond "test faster" on close to 0% of occasions and when the test and standard had the same speed, observers would perform at chance. However, when the test and standard had opposite directions of additional x-axis motion added, the observers responded almost always "faster" or always "slower" and so, when averaged between the two "different" conditions (Fig. 5c and its mirror image), performance was close to a d' of 0 (50% correct) in each case. These results show that, when observers were forced to use only one eye, they showed very different patterns of performance for the "same" and "different" conditions.
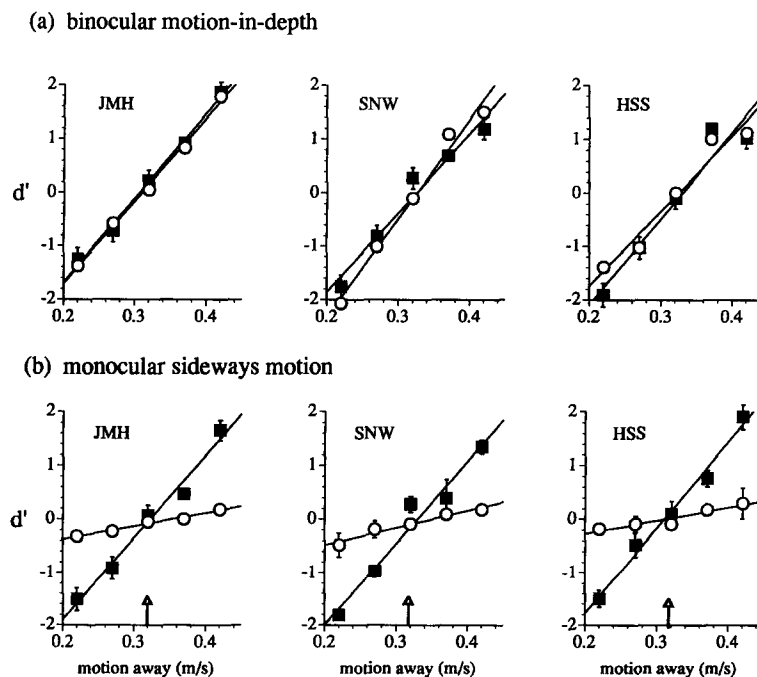

(a) binocular motion-in-depth

(b) monocular sideways motion

FIGURE 6. d' as a function of speed in the z-direction for the conditions where the test *and* standard had x-direction motion added or subtracted—the "same" condition—(■) and where one had x-direction motion added and the other had it subtracted—the "different" condition—(○). Results are shown for both the binocular motion-in-depth stimuli (a) and the monocular condition (b). For the monocular condition, we show both the monocular retinal speed (deg/sec) and the corresponding speed-in-depth (m/sec) that was present in the 3-D stimulus. Error bars show SEs calculated from at least three experimental runs of 100 trails. As discussed in the text, the monocular "different" condition (○, lower graphs) resulted in responses close to chance. However, the motion in depth "different" condition (■, upper graphs) showed very similar responses to the motion-in-depth "same" condition, suggesting that there is a binocular mechanism that combines the two monocular speeds to find 3-D speed (see text).

For the binocular motion-in-depth condition (upper graphs), observers gave very similar responses in both the "same" and "different" conditions. The psychometric functions were also very similar to those for the monocular "same" condition. This suggests that for these binocular conditions, observers did not simply use a monocular cue. If they had, the "different" condition should look as it did in the monocular condition. Instead, the results suggest that observers are using a combination of the two monocular cues to obtain the speed-in-depth of the moving dot.

It is also important to note that similar thresholds were found here and in Experiment I. Although not shown in Fig. 6, Weber fractions here were 0.14 for observer JMH, 0.2 for SNW and 0.22 for HSS (compared with 0.13 for JMH, 0.16 for SNW and 0.19 for HSS in Experiment I).

Beverley and Regan (1973) suggested that the direction of motion-in-depth was calculated from the *ratio* between the speeds of motion in each eye. Our results support the idea that a comparison or combination of the monocular speeds could also be used to extract the speed of motion-in-depth, but the data do not suggest the form of the combination. For instance, it would be possible to use the sum (or the average) of the modulus of each speed to discriminate which speed was faster. Another strategy might be to compare the signals from the two eyes and use the largest. Both these strategies would enable observers to solve the task, but not necessarily to know the 3-D speed and the 3-D direction of the motion. So, although the results suggest a binocular mechanism which takes the left and right eyes' monocular speeds as its input, the results do not prove that the visual system is able to calculate 3-D speed from a combination of the left and right eyes' monocular motion signals.

GENERAL DISCUSSION

The experiments presented here tested whether observers are able to judge the speed of motion along the z-axis (towards or away from the observer), and if so, what cues are used. We found that the speed of motion-in-depth can be discriminated as accurately as that of the equivalent 2-D motion (the motion in one stereo half-image) when the stimulus contains both stereo and monocular cues (and note that our analysis showed that observers used speed cues, rather than disparity or position cues). For a stimulus containing only disparity cues, speed-in-depth judgements were very poor, and were based at least partly on static disparity, rather than speed cues. This suggests that the rate-of-change of stereo disparity is a poor cue to the discrimination of speed-in-depth. Finally, we showed that individual monocular cues are not being used to solve the speed discrimination task and therefore that the cue to the discrimination of the speed of motion-in-depth is a combination of the two monocular speed cues.

Below, we discuss three aspects of our findings in this study. First, we consider the possible mechanisms that might be used for obtaining speed-in-depth, in the light of our data and the work of others. We suggest why monocular motion cues are important and whether static stereo cues are also required for speed-in-depth discrimination. Second, we discuss the use of dynamic random dot stereograms, and whether the use of short life-time dots is appropriate for this kind of experiment. Finally, we address whether we have found evidence for a mechanism designed to calculate the 3-D speed of moving objects.

Mechanisms for processing motion-in-depth

Recently, Cumming and Parker (1994) performed experiments showing that the *detection* of motion-in-depth can be based purely on disparity cues and does not require consistent monocular motion. The results presented here do not contradict this earlier result. The detection of motion-in-depth may be accomplished by responding to a change in static depth, and does not necessarily require a specific motion sensitive mechanism. However, the discrimination of speed requires motion information and our results suggest that the speed of motion-in-depth may be processed by a different mechanism from that described for detecting motion-in-depth. Consider the potential mechanisms shown in Fig. 7. The first mechanism (Fig. 7a) suggests that the speed of motion-in-depth is found from the rate of change of retinal disparity. Although this mechanism was shown to be adequate for the *detection* of motion-in-depth (Cumming & Parker, 1994) we have shown here that it is not adequate for accurately judging the speed of motion-in-depth. Our results favor either the second mechanism (Fig. 7b), in which the monocular speed signals from the two eyes are combined to find the speed-in-depth, or the third mechanism (Fig. 7c), in which although motion information is used to find the speed, consistent stereo disparity is also required.

Why might the speed of motion-in-depth be based on monocular motion cues, rather than disparity cues? Firstly, binocular disparity might not be used for calculating the speed of objects moving in depth because of the poor spatial and temporal resolution of stereopsis. Regan and Beverley (1973) showed that when a bar is moved back and forth fast enough in depth (with a temporal frequency of around 8 Hz), no motion-in-depth is seen. Thus the stereo system is relatively temporally "sluggish". It is also known that depth discrimination is very poor away from the fixation plane (Blakemore, 1970; McKee, Levi & Bowne 1990). If an object is moving fast enough that it spends only a short time near the fixation plane, the stereo signal could be very poor. Thus, is it perhaps unsurprising that poor speed discrimination is obtained based on stereo disparity cues alone.

A second reason why monocular motion cues are used might be due to the complexity of processing motion. Cumming and Parker (1994) argued that solving the stereo correspondence problem may require considerable processing and that it would be potentially wasteful to have to re-solve that problem if monocular
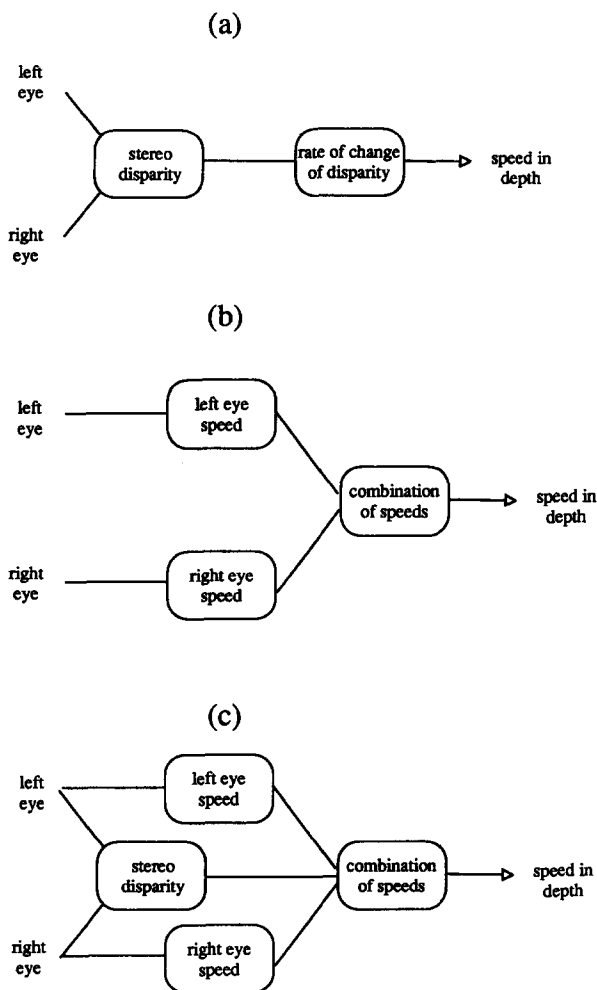
(a)

(b)

(c)

FIGURE 7. Potential mechanisms for finding the speed of motion-in-depth. (a) Information from the left and right eyes is combined to give stereo disparity. The rate of change of disparity is calculated to give the speed-in-depth. (b) Information from the left and right eyes is used independently to find the monocular speeds. Then, the speed information is combined to find the speed-in-depth. (c) A combination of the first two mechanisms. Stereo disparity is required to find the speed of motion-in-depth, as well as a comparison of monocular speeds.

information were combined at a later stage. However, it may be the case that the calculation of local speed is also a difficult problem. Indeed, most models for local speed calculation require complex processing in several states (e.g. Heeger, 1987; Grzywacz & Yuille, 1990). It might be the case that because obtaining speed is such a difficult computation for the brain, it is done just once using monocular motion and is later interpreted in 3-D by comparison processes. However, as Cumming and Parker pointed out, because stereo correspondence is such a hard problem, it may also be performed only once and thus it may be necessary to have a consistent stereo signal before a comparison of monocular motion signals can take place. Our experiments did not test whether stereo correspondence is required in order to be able to compare monocular speeds.

There is some evidence which suggests that consistent stereo signals may be necessary for there to be any comparison of monocular speed signals to give the speed of motion-in-depth. Cumming and Parker (1994) found

that motion cues without consistent disparity could not be used to *detect* motion-in-depth. The stimuli they used to test this relied on the fact that the stereo system has relatively poor resolution in both space and time (Tyler, 1983). In one experiment, Cumming and Parker essentially removed the disparity cue by using stimuli that contained a fine-scale depth-defined square-wave grating that was beyond the spatial range of stereopsis (the corrugations in depth could not be resolved). In a different experiment, they used a stimulus where the motion-in-depth occurred at a temporal frequency too high to be detected by the stereo system. Although motion could be seen monocularly, in neither of these stimuli could observers detect motion-in-depth. Presumably, if motion-in-depth could not be detected, then it would not be possible to judge its speed. Thus, these experiments suggest that stereo must be present before speed-in-depth can be extracted.

Another series of studies where monocular motion is observed but no motion-in-depth is detected are those of Erkelens et al. (Erkelens & Collewijn, 1985a, b; Regan, Erkelens & Collewijn, 1986). They used a range of stereograms in which the whole of the left and right eyes' fields were moved in opposite directions. For a single dot stimulus, motion-in-depth thresholds were found to be very high and for other stimuli, no motion-in-depth was detected at all. However, eye movement recordings showed that there was motion on each retina. Again, we have a stimulus in which there are monocular motion cues, but they cannot be used to obtain the speed of motion-in-depth because that motion cannot even be detected. We suggest two possible reasons why observers might not be able to use the available monocular motion cues to find the speed-in-depth. First, accurate speed discrimination relies on relative motion. If the monocular motion units that are combined to feed the speed-in-depth mechanism were sensitive to relative motion (as suggested by Regan, 1986; Regan et al. 1986), then they would not give an appropriate response when there was no stationary reference. Thus, it would not be possible to obtain an accurate measure of the speed-in-depth.

Second, speed-in-depth may require consistent relative disparities to be present. This argument is similar to the one we used when discussing the Cumming and Parker (1994) experiments. In that study, there were no consistent relative disparities in the stimulus. Although absolute disparities were present in the Erkelens et al. stimulus, as shown by the eye movement recordings, there were no relative disparities in the stimulus, to which the stereo system is particularly sensitive (Westheimer, 1979). Hence, if relative disparities are required before the speed-in-depth can be extracted, we would not expect to be able to obtain the speed-in-depth for such stimuli.

Use of dynamic random dot stereograms

In Experiment II, we used dynamic random dot stereograms so that consistent monocular motion cues were removed from the display. It could be argued that

the disparity mechanism may perform poorly when presented with stimuli containing such short lifetime elements. Indeed, for each new frame of the stereogram, there is in principle a new corresponding problem. If this were problematic for stereo processing, speed discrimination for dynamic random dot stereograms might be poor because elements were transient, rather than because the stimulus did not contain monocular motion cues. We cannot be certain that this hypothesis is wrong because we were unable to design a stimulus containing longer lifetime dots *and* no monocular motion cues. However, there is evidence suggesting that static and dynamic stereo stimuli provide an equivalent stereo signal in both stationary stereo displays and when stereo-defined objects move. As part of a recent study (Harris & Watamaniuk, 1994b) we measured steroacuity thresholds for a patch standing out in depth from a standard random dot stereogram (SRDS) and from a dynamic random dot stereogram (DRDS). For two of our three observers, thresholds for detection of the patch were very similar for the two stimuli (JMH: 0.12 min arc for SRDS, 0.12 min arc for DRDS; SNW: 0.12 min arc for SRDS, 0.18 min arc for DRDS), while the threshold for dynamic stereo was three times that for standard stereo for our third observer (HSS: 0.27 min arc for SRDS; 0.83 min arc for DRDS). Note that the range of stereo disparity used in the motion-in-depth experiment was +14 to −14 min arc, many times above threshold for all observers.

Another piece of evidence that dynamic stimuli are equivalent to standard stereo stimuli comes from Stevenson, Cormack and Schor (1994) who showed that vergence can be driven by dynamic stereo (as it can for standard stereo stimuli, see e.g. Erkelens & Collewijn, 1985a, b).

For stereo-defined regions moving in the fronto-parallel plane, Patterson, Bowd, Phinney, Pohndorf, Barton-Howard and Angilletta (1994) showed that motion after-effects can be obtained using both standard stereograms and dynamic stereograms.

Finally, Cumming and Parker (1994) found that motion-in-depth can be detected equally well using dynamic or constant dot lifetime random dot stereograms. Any system that differentiated the disparity signal to give a speed signal would be limited by the noise in the disparity detection mechanism. Since motion-in-depth is detected equally well for static or dynamic stereograms, the signal to be differentiated should be equally noisy, or conversely equally good. The Cumming and Parker result agrees with the subjective impression of observers in our motion-in-depth experiment. Observers always saw the motion-in-depth very clearly, whether presented using a standard stereogram or a dynamic stereogram.

We found no studies in the literature in which the stereo signal from dynamic random dot stereograms was shown to be significantly poorer than that from a standard random dot stereogram. Thus, the disparity mechanism appears to respond equally well to standard stereograms and dynamic stereograms. Poor speed-in-

depth discrimination with the dynamic stereogram cannot be accounted for by the relative strengths of the stereo signal from the standard and dynamic stereograms.

Does the visual system calculate 3-D speed from a combination of monocular motion signals?

Our experiments suggest the existence of a binocular mechanism for discriminating the speed of motion-in-depth based on a comparison of the monocular speed cues from the two eyes. However, we do not have sufficient evidence to assert that the mechanism calculates 3-D speed. In the experiments presented here, we asked observers to discriminate between speeds along a single 3-D direction, namely that of motion in the z-direction, directly towards or away from the observer. In order to test whether 3-D speed was calculated from a combination of monocular motion cues, it would be necessary to measure speed discrimination for different 3-D directions. If the visual system calculated 3-D speed from the monocular motion signals, it would be expected that motions along different 3-D trajectories would appear to have the same speed, even though the monocular retinal motions could be very different.

It is possible that we have found evidence of a binocular mechanism that combines speeds from the two eyes but for a purpose other than finding 3-D speed. One potential mechanism might sum the motion energy (Adelson & Bergen, 1985, Watson & Ahumada, 1985) from the two eyes in order to improve the 2-D signal, rather than to find 3-D speed. If such a mechanism used motion energy but was insensitive to direction (in other words, if it did not have access to the *sign* of the velocity), it would not be possible to find 3-D speed but it would be possible to solve our experimental task. In further experiments, we plan to test whether the binocular mechanism found here contributes to a 3-D representation of speed or whether it has a different purpose.

REFERENCES

Adelson, E. H. & Bergen, J. R. (1985). Spatio-temporal energy models for the perception of motion. *Journal of the Optical Society of America A, 2,* 284–299.

Beverley, K. & Regan, D. R. (1973). Evidence for the existence of neural mechanisms selectively sensitive to the direction of movement in space. *Journal of Physiology, London, 235,* 17–29.

Beverley, K. & Regan, D. R. (1974). Visual sensitivity to disparity pulses: evidence for direction selectivity. *Vision Research, 14,* 357–361.

Beverley, K. & Regan, D. R. (1975). The relation between discrimination and sensitivity in the perception of motion in depth. *Journal of Physiology, London, 249,* 387–398.

Blakemore, C. (1970). The range and scope of binocular depth discrimination in man. *Journal of Physiology, London, 211,* 599–622.

Cumming, B. G. & Parker, A. J. (1994). Binocular mechanisms for detecting motion-in-depth. *Vision Research, 34,* 483–496.

Erkelens, C. J. & Collewijn, H. (1985a). Eye movements and stereopsis during dichoptic viewing of moving random-dot stereograms. *Vision Research, 25,* 1689–1700.

Erkelens, C. J. & Collewijn, H. (1985b). Motion perception during dichoptic viewing of moving random-dot stereograms. *Vision Research, 25,* 583–588.

Finney, D. J. (1971). *Probit analysis* (3rd edn). Cambridge: Cambridge University Press.

Grzywacz, N. M. & Yuille, A. L. (1990). A model for the estimate of local image velocity by cells in the visual cortex. *Proceedings of the Royal Society of London, B, 239*, 129–161.

Harris, J. M. & Watamaniuk, S. N. J. (1994a). Speed discrimination of binocular motion in depth. *Investigative Ophthalmology and Visual Science, 35*, 1986.

Harris, J. M. & Watamaniuk, S. N. J. (1994b). Speed discrimination of motion defined by binocular disparity. *Perception. 23*, 60.

Heeger, D. J. (1987). A model for the extraction of image flow. *Journal of the Optical Society of America, A, 4*, 1455–1471.

Julesz, B. (1971). *Foundations of cyclopean perception*. Chicago: University of Chicago Press.

McKee, S. P. (1981). A local mechanism for differential velocity detection. *Vision Research, 21*, 491–500.

McKee, S. P., Levi, D. M. & Bowne, S. F. (1990). The imprecision of stereopsis. *Vision Research, 30*, 1763–1779.

Orban, G. A., de Wolf, J. & Maes, H. (1984). Factors influencing velocity coding in the human visual system. *Vision Research, 24*, 33–39.

Patterson, R., Bowd, C., Phinney, R., Pohndorf, R., Barton-Howard, W. J. & Angilletta, M. (1994). Properties of the stereoscopic (cyclopean) motion aftereffect. *Vision Research, 34*, 1139–1147.

Regan, D. (1986). Visual processing of four kinds of relative motion. *Vision Research, 26*, 127–145.

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

Regan, D. & Beverley, K. (1973a). Some dynamic features of depth perception. *Vision Research, 13*, 2369–2378.

Regan, D. & Beverley, K. (1973b). The dissociation of sideways movements from movements in depth: psychophysics. *Vision Research, 13*, 2403–2415.

Regan, D. & Beverley, K. (1978a). Illusory motion in depth: aftereffect of adaptation to changing size. *Vision Research, 18*, 209–212.

Regan, D. & Beverley, K. (1978b). Looming detectors in the human visual pathway. *Vision Research, 18*, 415–421.

Regan, D., Erkelens, C. J. & Collewijn, H. (1986). Necessary conditions for the perception of motion in depth. *Investigative Ophthalmology and Visual Science, 27*, 584–597.

Stevenson, S. B., Cormack, L. K. & Schor, C. M. (1994). The effect of stimulus contrast and interocular correlation on disparity vergence. *Vision Research, 34*, 383–396.

Tyler, C. W. (1971). Stereoscopic depth movement: two eyes less sensitive than one. *Science, New York, 174*, 958–961.

Tyler, C. W. (1983). Sensory processing of binocular disparity. In Schor, C. M. & Cieffreda, K. J. (Eds), *Vergence eye movements: basic and clinical aspects*. Boston, Mass.: Butterworth.

Tyler, C. W. & Foley, J. M. (1974). Stereomovement suppression for transient disparity changes. *Perception, 3*, 287–296.

Watson, A. B. & Ahumada, A. J. (1985). Model of human visual motion sensing. *Journal of the Optical Society of America, A, 2*, 322–342.

Westheimer, G. (1979). Cooperative neural processes involved in stereoscopic acuity. *Experimental Brain Research, 36*, 585.

Wheatstone, C. (1838). Conditions to the physiology of vision—1. On some remarkable, and hitherto unobserved, phenomena of binocular vision. *Philosophical Transactions of the Royal Society of London, B, 128*, 371–394.