

## No evidence for sequential effects of the interaction of stereo and motion cues in judgements of perceived shape

Rebecca A. Champion<sup>a</sup>, Eli Brenner<sup>b,\*</sup>, Pascal Mamassian<sup>a</sup>, David R. Simmons<sup>a</sup>

<sup>a</sup> Department of Psychology, University of Glasgow, 58 Hillhead Street, Glasgow G12 8QB, Scotland, UK

<sup>b</sup> Department of Neuroscience, ErasmusMC, postbus 1738, 3000 DR Rotterdam, The Netherlands

Received 10 December 2002; received in revised form 21 July 2003

### Abstract

The interaction of the depth cues of binocular disparity and motion parallax could potentially be used by the visual system to recover an estimate of the viewing distance. The present study investigated whether an interaction of stereo and motion has effects that persist over time to influence the perception of shape when the motion information is removed. Static stereoscopic ellipsoids were presented following the presentation of rotating stereoscopic ellipsoids, which were located either at the same or a different viewing distance. It was predicted that shape judgements for static stimuli would be better after presentation of a rotating stimulus at the same viewing distance, than after presentation of one at a different viewing distance. No such difference was found. It was concluded that an interaction between stereo and motion depth cues does not influence the perception of subsequently presented static objects.

© 2003 Elsevier Ltd. All rights reserved.

**Keywords:** Depth; Three-dimensional shape; Stereopsis; Motion parallax; Cue interaction

### 1. Introduction

To recover information about Euclidean shape from binocular disparity, knowledge of the viewing distance and inter-ocular separation are required. In the case of shape from motion, information about eye rotation and ego motion are required. However, when binocular disparity and motion parallax are presented together, veridical information about depth and distance can be recovered without explicit knowledge of these scaling parameters (Richards, 1985).

Although some evidence in support of this scheme has been provided (Econopouly & Landy, 1995; Johnston, Cumming, & Landy, 1994) more recent evidence suggests that such an interaction does not occur (Bradshaw, Parton, & Glennerster, 2000; Brenner & Landy, 1999; Brenner & van Damme, 1999; Landy & Brenner, 2001). Brenner and van Damme showed that any estimate of viewing distance that may have been recovered from such an interaction is not transferred to

other attributes than shape (i.e. there was no change in perceived size and distance). Brenner and Landy subsequently demonstrated that any estimate of viewing distance that may have been recovered from such an interaction is not transferred across space to influence the shape of an object at a different viewing distance. Here we investigate the last remaining possibility: that an estimate of viewing distance is recovered and transferred *across time* to influence the shape of an object that is subsequently presented at the same location.

In the present study, participants were presented with ellipsoids, defined either by both motion and stereo cues (rotating stimuli), or stereo cues alone (static stimuli). Stimuli of each type were presented alternately. In experiment 1, the initial size and shape of the ellipsoid was varied at random and the participant's task was to set the ellipsoid to the size and shape of a tennis ball. In experiment 2, the size of the ellipsoid was fixed at that of a tennis ball. This was done to eliminate the possibility that the random change in size induces an apparent change in viewing distance even when such a change was not simulated. Hence, in experiment 2, the participant's task was simply to set the shape of the ellipsoid to a sphere. Stimuli were presented at two different viewing

\* Corresponding author. Tel.: +31-10-4087569; fax: +31-10-4087462.

E-mail address: [e.brenner@erasmusmc.nl](mailto:e.brenner@erasmusmc.nl) (E. Brenner).

distances, 40 and 80 cm (the viewing distance will henceforth be termed the target distance, as it refers to the distance to the simulated target rather than the actual distance to the screen or to where the subject is looking).

We were interested in looking at the effect of either changing or not changing the target distance between rotating and subsequent static trials, on the perceived shape of the static stimuli. This factor was termed the ‘temporal structure’, and consisted of three different conditions. Blocks of trials had either randomly changing or constant target distance. Within the random blocks, there was either no change, or a change, in target distance between the presentation of each stimulus and the previous stimulus. These conditions were termed the *random no-change* and *random change* conditions respectively. Within the constant target distance blocks, there were no changes in target distance between trials. This condition was termed the *blocked no-change* condition (only settings made in the second half of these blocks were analysed to make sure that each target was preceded by many targets at the same distance).

Our hypothesis was that the visual system recovers a veridical estimate of the viewing distance during the presentation of a rotating stimulus, but that this distance estimate can only be used to interpret the disparities of a subsequent static ellipsoid if there was *no change* of target distance. Hence, we examined whether the shape judgements of the static stimuli would become more veridical following the presentation of a rotating stimulus when there was *no change*, compared to when there was a *change* in target distance.

## 2. Methods

### 2.1. Equipment

Images were presented with a Silicon Graphics Onyx RealityEngine on a high-resolution monitor (120 Hz; horizontal size: 39.2 cm, 815 pixels; vertical size: 29.3 cm, 611 pixels; spatial resolution refined with anti-aliasing techniques). Subjects sat with their head in a chin-rest at 60 cm from the screen. The images were viewed through liquid crystal shutter spectacles that were synchronised with the refresh rate of the monitor. Alternate images were presented to the left and right eye, so that each eye received a new image every 16.7 ms (60 Hz). Red stimuli (and an additional red filter in front of the monitor screen) were used because the shutter spectacles have the least cross-talk at long wavelengths.

The ellipsoid was rendered in perspective projection, taking the individual’s inter-ocular distance into consideration. Therefore both the subject’s ocular convergence when fixating the ellipsoid and the images on his

or her retinas were appropriate for an ellipsoid at the simulated distance.

### 2.2. Stimuli

The stimulus was a computer-simulated opaque ellipsoid of which only the surface texture was visible. The axes of the ellipsoid were such that, when stationary, two axes of equal length were in the frontoparallel plane (the width and height dimensions), and the third axis, which could be longer, shorter or equal in length to the other axes, was along the line of sight (depth dimension). The texture on the ellipsoid’s surface consisted of 3000 randomly oriented triangles, about half of which were visible. The triangles were ‘painted’ onto the surface. When the ellipsoid was spherical the triangles were equilateral, with sides of 6% of the radius, and with randomly chosen positions and orientations on the surface. When the ellipsoid was not spherical, the triangles were stretched along the long axis of the ellipsoid.

The ellipsoid’s simulated distance was either 40 or 80 cm. The ellipsoids were either static or rotating. In the rotating condition, the ellipsoids rotated sinusoidally up and down around a horizontal axis (0.25 Hz,  $\pm 15^\circ$ ). The axis of rotation passed through the centre of the ellipsoid and was orthogonal to the line of sight.

Care was taken to ensure that no structures other than the simulated ellipsoids were ever visible. The tabletop and wall were covered with black cloth to reduce reflection, and the stimuli were red and quite dim.

As the images were rendered in the appropriate perspective for each eye, the stimulus contained texture cues as well as binocular disparities. These cues were always consistent with the simulated shape. Thus, texture, motion parallax, binocular disparity, and the vergence required to fixate any point on the object, were all consistent with an ellipsoid at the simulated distance. The only inconsistencies in the stimulus were a conflict with accommodation, the absence of a blur gradient, and the absence of motion parallax during any unintended head movement (we used a chin-rest rather than a bite-board). Shading provided no useful information (surfaces were rendered with uniform illumination).

### 2.3. Procedure

In experiment 1, the participant’s task was to set the size and shape of the simulated ellipsoids to match a tennis ball (radius 3.3 cm). In experiment 2, the participant’s task was only to set the shape of the ellipsoid to a sphere. During the experiments, observers held a real tennis ball in their left hand and the computer mouse in their right hand. Observers were encouraged to look at the tennis ball before, but not during, each session. In experiment 1, they adjusted the simulated ellipsoid’s width and depth by moving the computer mouse. Hor-

horizontal mouse movements changed the width of the simulated ellipsoid. The radius could vary between 1 and 10 cm. Vertical mouse movements changed the simulated depth of the ellipsoid. The depth could vary from 1/3 of the width to 3 times the width. The initial simulated width and depth of each ellipsoid were determined at random for each trial. In experiment 2, the simulated width was fixed at 3.3 cm, but the simulated depth was determined at random. Participants only performed vertical mouse movements to adjust the simulated depth. Observers indicated when they were satisfied with their settings by pressing the mouse button. The next target appeared immediately.

Stimuli were presented in blocks of 30 trials. In all blocks, static and rotating ellipsoids were alternated. Blocks had either randomly changing target distance, or constant target distance. In the random target distance blocks, the distance was chosen at random for each trial to be either 40 or 80 cm. In the constant target distance blocks, target distance was kept constant at either the near target distance or far target distance within the block. Observers each completed two sessions consisting of three blocks each. In each session the random block was presented first, followed by the two constant blocks, with the order of the constant blocks being counter balanced across the sessions. Each session took between 30 and 45 min to complete.

#### 2.4. Participants

Five participants took part. Two were authors, and the other three were naive as to the purpose of the experiment. All had normal binocular vision. Another participant was initially tested, but inspection of his data revealed that the variability of his shape judgements were more than four standard deviations higher than the mean variability of the other five participants. Hence, his results were excluded from the analysis.

#### 2.5. Analysis

##### 2.5.1. Shape error

For each setting, the shape error (SE) was calculated. This is a measure of how the set shape deviated from that of a sphere and is independent of size. The formula used for this measure was:

$$\text{Shape Error} = \frac{\text{depth} - \text{width}}{\text{depth} + \text{width}} \times 100$$

Hence, a shape error of 0% represents veridical settings. Positive shape errors represent setting of the depth to be greater than the width, or setting the shape to be 'stretched' along the depth dimension. Negative errors represent setting of the depth to be less than the width, or setting the shape to be 'squashed' along the depth

dimension. Note: In experiment 2 the width was fixed at 3.3 cm.

##### 2.5.2. Width error

For experiment 1, a width error (WE) was also calculated. The width error represents the degree to which the set width deviated from that of a tennis ball (3.3 cm), and was calculated as:

$$\text{Width Error} = \frac{\text{width} - 3.3}{\text{width} + 3.3} \times 100$$

Hence, a width error of 0% represents a veridical setting. Positive width errors represent setting of the width to be too big and negative width errors represent setting of the width to be too small.

### 3. Results

Figs. 1 and 2 show the shape errors found in experiments 1 and 2 respectively. In these figures, shape errors, averaged across participants, are shown for the static and rotating stimuli in the three temporal structure conditions, at both target distances. Both figures demonstrate a large difference between the effect of target distance on the static and rotating settings. For the static stimuli, the shape settings were significantly different for the two target distances. The pattern of results is consistent with participants under-compensating for the difference in target distance when scaling the disparities, as has been found in previous studies (e.g. Johnston, 1991). For the rotating stimuli there was no effect of target distance. The figures also demonstrate that there

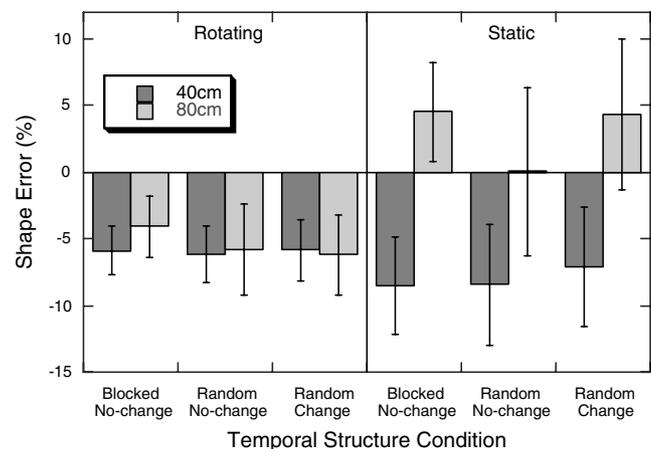


Fig. 1. Mean shape errors for experiment 1. Individual bars represent shape errors, averaged across participants, for static and rotating stimuli. Data are split into three temporal structure conditions according to whether there was a *change*, or *no change* in target distance relative to the previous target, and also according to whether target distance was *blocked* or *randomly changing* within the block. Data are also grouped by target distances. Error bars represent  $\pm 1$  standard error across participants.

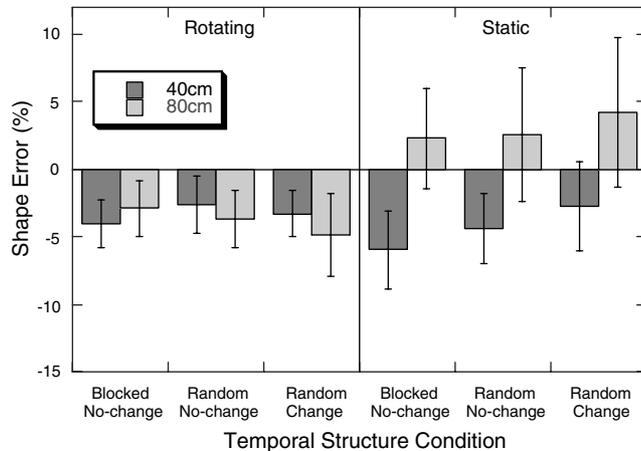


Fig. 2. Mean shape errors for experiment 2. Individual bars represent shape errors, averaged across participants, for static and rotating stimuli, in the three temporal structure conditions at both target distances. Error bars represent  $\pm 1$  standard error across participants.

was little effect of the different temporal structure conditions on static settings.

The means of the static settings from each experiment were analysed with a 2-factor, repeated-measures ANOVA (rotating settings were not included as they were of less interest). The effect of distance was found to be significant in experiment 1 ( $F_{1,4} = 15.8$ ,  $p < 0.05$ ) though it was not significant in experiment 2 ( $F_{1,4} = 3.2$ , n.s.). This difference in the effect of distance in the two experiments reflects the fact that there was a greater discrepancy in average shape error between the two target distances in experiment 1 compared to experiment 2. This may be due to the subjects having had more practice at the task in experiment 2, or alternatively to the use of the fixed width as a cue to distance. Neither the effect of temporal structure (Exp 1:  $F_{2,8} = 1.1$ , n.s.; Exp 2:  $F_{2,8} = 3.1$ , n.s.) nor the interaction of distance and temporal structure (Exp 1:  $F_{2,8} = 1.5$ , n.s.; Exp 2:  $F_{2,8} = 0.3$ , n.s.) were found to be significant in either experiment. The settings of the rotating stimuli were not veridical. However, they were not always biased in the same direction as the static stimuli, so the bias is unlikely to be related to misjudgment of the viewing distance.

We also looked at whether there was any effect of temporal structure on width settings in experiment 1. Fig. 3 shows the width errors averaged across participants for the static and rotating stimuli in the three temporal structure conditions, at both target distances. This figure shows clearly that the change in target distance has the same effect on both static and rotating width errors. As with the shape errors for the static stimuli, this pattern of errors is consistent with participants under-compensating for the different target distances when scaling the retinal size. A 2-factor, repeated-measures ANOVA found a significant effect of distance ( $F_{1,4} = 10.3$ ,  $p < 0.05$ ), but no effect of temporal

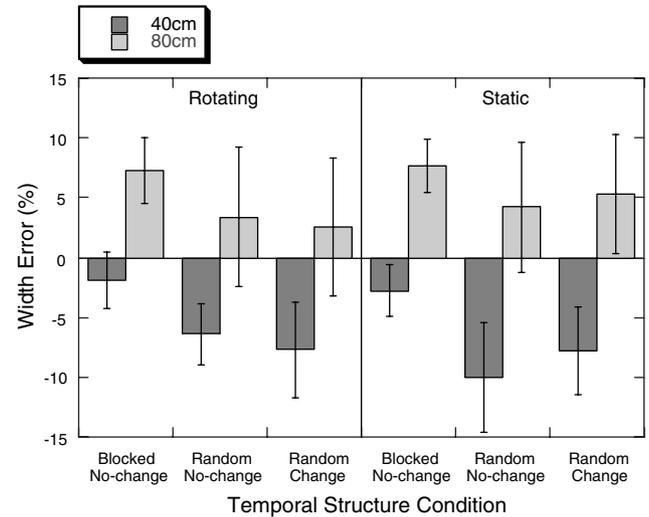


Fig. 3. Mean width errors for experiment 1. Individual bars represent width errors, averaged across participants, for static and rotating stimuli, in the three temporal structure conditions at both target distances. Error bars represent  $\pm 1$  standard error across participants.

structure ( $F_{2,8} = 2.04$ , n.s.) and no interaction between temporal structure and distance ( $F_{2,8} = 0.38$ , n.s.).

#### 4. Discussion

The aim of this study was to investigate the possibility that an estimate of viewing distance, recovered through the interaction of stereo and motion, is linked to a specific location in space, but can be transferred *over time* to a new object at the same location. We predicted that if so we would find an improvement in shape settings for static stimuli following the presentation of a rotating stimulus when there was *no change* in target distance, but not when there was a *change* of target distance. This prediction was not supported by two experiments.

These findings, therefore, support the findings of Brenner and van Damme (1999) and Brenner and Landy (1999) by providing further evidence against the proposal that an interaction between stereo and motion depth cues allows the visual system to recover a more veridical estimate of viewing distance.

#### Acknowledgements

R.A. Champion is supported by a studentship from the Economic and Social Research Council (ESCR). Also this work is supported in part by the European Commission Research Training Network "Perception for Recognition and Action" (HPRN-CT-2002-00226).

## References

- Bradshaw, M. F., Parton, A. D., & Glennerster, A. (2000). The task-dependent use of binocular disparity and motion parallax information. *Vision Research*, *40*, 3725–3734.
- Brenner, E., & Landy, M. S. (1999). Interaction between the perceived shape of two objects. *Vision Research*, *39*, 3834–3848.
- Brenner, E., & van Damme, W. (1999). Perceived distance, shape and size. *Vision Research*, *39*, 975–986.
- Econopouly, J. C., & Landy, M. S. (1995). Stereo and motion combined rescale stereo. *Investigative Ophthalmology and Visual Science*, *36*, S665.
- Johnston, E. B. (1991). Systematic distortions of shape from stereopsis. *Vision Research*, *31*, 1351–1360.
- Johnston, E. B., Cumming, B. G., & Landy, M. S. (1994). Integration of stereopsis and motion shape cues. *Vision Research*, *34*, 2259–2275.
- Landy, M. S., & Brenner, E. (2001). Motion-disparity interaction and the scaling of stereoscopic disparity. In M. Jenkin & L. Harris (Eds.), *Vision and attention* (pp. 129–150). New York: Springer-Verlag.
- Richards, W. (1985). Structure from stereo and motion. *Journal of the Optical Society of America*, *2*, 343–349.