PERGAMON

Vision Research 39 (1999) 3834–3848

Vision Research

www.elsevier.com/locate/visres

Interaction between the perceived shape of two objects

Eli Brenner [a,*], Michael S. Landy [b]

[a] Vakgroep Fysiologie, Erasmus Universiteit, Postbus 1738, 3000 DR Rotterdam, The Netherlands

[b] Department of Psychology and Center for Neural Science, New York University, 6 Washington Place, New York, NY 10003-6634, USA

Received 8 June 1998; received in revised form 30 October 1998

Abstract

The difference between the way in which binocular disparity scales with viewing distance and the way in which motion parallax scales with viewing distance introduces a potential indirect cue for viewing distance: the viewing distance is the only distance at which disparity and motion specify the same depth. The present study examines whether this information is used. Two simulated ellipsoids were presented on a computer screen in complete darkness. The two ellipsoids were 6° to the left and right of straight ahead. Subjects set the width and depth of each ellipsoid to match a tennis ball, and set the distance of the one on the right to half that of the one on the left. The distance of the left ellipsoid varied between trials. On half of the trials it was static. On the other half it was rotating up and down around its frontal horizontal axis. Rotating the left ellipsoid influenced its set depth: rotating ellipsoids were set to be much more spherical. There was no influence on the set depth of the other ellipsoid, or on the set width of either. The set distance of the right ellipsoid was also unaffected. We conclude that subjects do not combine binocular disparity and motion parallax to obtain more veridical information about viewing distance. © 1999 Elsevier Science Ltd. All rights reserved.

Keywords: Depth; Stereopsis; Motion parallax; Cue interaction

1. Introduction

When an isolated object is presented in an otherwise completely dark environment, its distance is often misjudged. Under such conditions, depth[1] judgements that are based solely on binocular disparities are also often incorrect (Johnston, 1991). Such errors co-vary with errors in judgements of width and height in a manner that suggests that an incorrect estimate of the viewing distance is responsible (van Damme & Brenner, 1997; Brenner & van Damme, 1999). When more than one object is presented, the spatial relationships between the objects are also misjudged in a manner that suggests that it is the estimate of viewing distance that is incorrect (Foley, 1980, 1985).

Under some conditions, rotating an isolated object results in much more veridical judgements of its shape (Johnston, Cumming & Landy, 1994). Presumably this is because misjudging the distance does not influence the shape judged from motion parallax as it does the shape judged from relative disparities, so that combining the two provides a better estimate of shape than disparities do on their own. Subjects may even give less weight to the binocular information when they know it to be less reliable than that from motion parallax (Gillam, 1968; Young, Landy & Maloney, 1993; Landy, Maloney, Johnston & Young, 1995). However, there is even more to be gained by combining the two cues.

The different ways in which misjudging the distance affects depth from binocular disparity and depth from motion parallax makes it possible to obtain a better estimate of the distance to the object (Richards, 1985; Johnston et al., 1994; Frisby, Buckley, Wishart, Porrill, Gårding & Mayhew, 1995). The actual viewing distance is the only distance for which the two cues indicate the same depth (see Fig. 1). Here, we examine whether observers use this source of information.

* Corresponding author. Tel.: +31-10-4087569; fax: +31-10-4089457.

E-mail address: brenner@fys.fgg.eur.nl (E. Brenner)

[1] *Depth* is used to refer to an object's dimension along the line of sight, while *distance* is used to indicate how far it is from the observer.

PII: S0042-6989(99)00105-4

In a study with two cylinders located at the same distance from the observer (Econopouly & Landy, 1995), rotating one of the cylinders usually resulted in more veridical judgements of shape; not only for the moving cylinder, but also for the second, static one. This finding supports the notion that the cues are combined to obtain a better estimate of distance, and that this refined distance estimate is used to interpret the disparities of other, static objects in the scene. In contrast, in a study with a single ellipsoid, rotating the ellipsoid resulted in more veridical judgements of its shape, but not of its size or position (Brenner & van Damme, 1999). Judgements of size and position should also have improved if a better estimate of distance had become available, so this study suggests that the estimate of distance was not refined, but that subjects ignored the binocular disparities when motion parallax provided more reliable information about depth (see Rogers & Collett, 1989; Tittle & Braunstein, 1993; Johnston et al., 1994; Tittle, Todd, Perotti & Norman, 1995 for evidence that binocular disparities are certainly not always ignored in the presence of depth information from motion parallax).

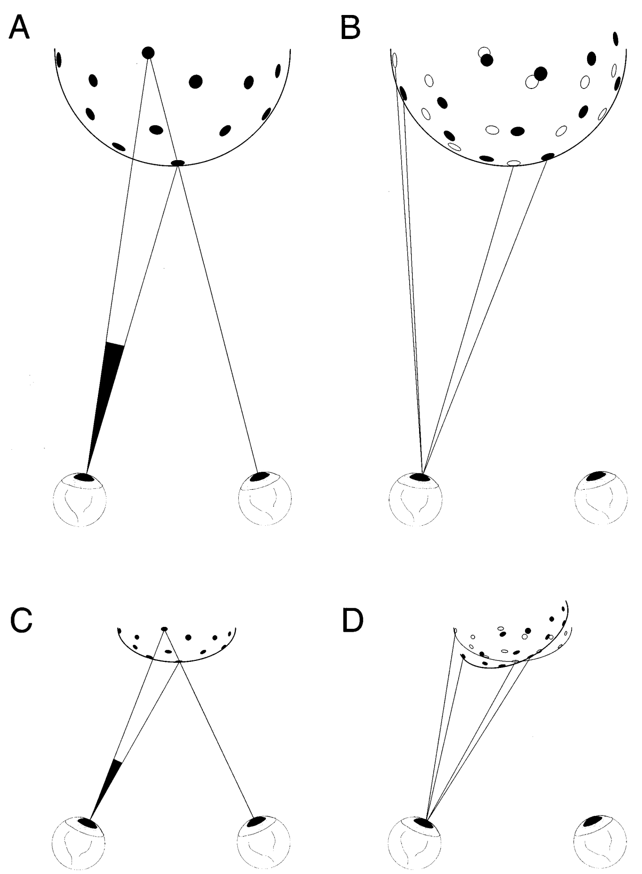

Fig. 1. Schematic diagram illustrating how combining depth cues could help determine an ellipsoid's distance from the observer. (A) Top view of two eyes and a sphere. The relative retinal disparity between two selected points on the surface of the sphere is indicated. For simplicity the points have been chosen so that they are aligned (horizontally) in the right eye. (B) Angular displacement of two dots (in the left eye) as the sphere rotates. (C) An object that would give rise to the same relative disparity between the two dots and would have the same angular extent (i.e. that would produce the same retinal images) at a smaller distance. Notice the different relationship between width and depth. (D) Rotating the nearer object gives a very different pattern of angular velocities than in (B). Thus, for every combination of relative disparities and angular velocities there is only one object distance and shape that is consistent with both cues.

The two findings would be reconciled if we could show that the improvement is restricted to judgements of shape. Visual processing takes place within many relatively independent pathways (Marr, 1982; Livingstone & Hubel, 1988; Zeki & Shipp, 1988; Glennerster, Rogers & Bradshaw, 1996; Tittle & Perotti, 1997). Thus, information gained within one pathway need not influence processing within other pathways. Consequently, adding information about the shape of one object could cause the shapes of other objects to be perceived more veridically, without their perceived size or position being affected. Alternatively, the estimate of viewing distance may not have improved at all in Econopouly and Landy's study. Since the two objects were always at the same distance, subjects may have been directly comparing the disparity of the static object to that of the moving one (taking the perceived shape into consideration).

In the present study we distinguish between these two alternatives by having subjects make settings for two objects at different distances. When the objects are at different distances, it is no use comparing disparities directly, because the relationship between disparity and depth is not the same at different distances. However, we do expect an improvement in the judged distance of one of the objects to transfer to the other object when the subject shifts his or her gaze between them, because subjects have been shown not to lose their sense of distance across such changes in fixation (Brenner & van Damme, 1998) and to be capable of making accurate binocular depth matches under such conditions (Glennerster et al., 1996). The two objects were presented at eye height in an otherwise dark environment; conditions under which distance is known to be misjudged. Subjects were restrained by a chin rest, but were free to look wherever they wanted.

2. Methods

In the main experiment, subjects set the size and shape of two binocular simulations of randomly textured ellipsoids to match a tennis ball. In addition, they set the simulated distance of the ellipsoid on the right to half that of the one on the left. The ellipsoid on the left could either be moving or static, whereas that on the right was always static. The same subjects also performed the same task in two additional experiments.

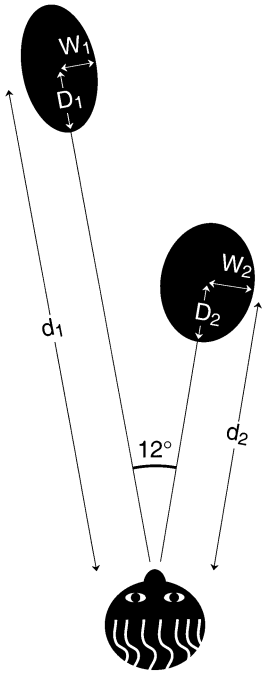

Fig. 2. Schematic view of the simulation as seen from above. Subjects could change the simulated widths (W_1,W_2) and depths (D_1,D_2) of both ellipsoids, and the distance of the right ellipsoid (d_2), by moving the computer mouse. The distance of the left ellipsoid (d_1) was chosen at random for each trial.

The first was identical to the main experiment except that the ellipsoid on the left was always moving, and the one on the right was to be set to the same distance as that on the left. The second additional experiment was identical to the first except that subjects were asked to set the size of the right ellipsoid to twice the size of a tennis ball.

2.1. Equipment

A schematic representation of the simulated scene is shown in Fig. 2. 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 80 cm from the screen. The images were viewed through liquid crystal shutter spectacles which 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).

Each image was presented in accordance with the way in which an ellipsoid would be seen from the position of the eye for which it was intended (taking the individual's inter-ocular distance into consideration), so that both the subject's ocular convergence when fixating the ellipsoid and the images on his or her retinae were appropriate for an ellipsoid at the simulated distance. Red stimuli (and additional red filters in front of the spectacles) were used because the shutter spectacles have the least cross-talk at long wavelengths.

2.2. Stimuli

The stimuli were computer-simulated opaque ellipsoids, of which only the surface texture was visible (see Fig. 3). The texture on the ellipsoid's surface consisted of 1800 randomly oriented triangles, about half of which were visible. The triangles were 'painted' onto the surface so that they 'stretched' (in the simulation) when the ellipsoid's shape was changed. When the ellipsoid was spherical the triangles were equilateral (with sides of 6% of the radius) and were distributed and oriented at random on the surface.

The simulated distance of the ellipsoid on the left was between 80 and 200 cm from the subject's eyes. The value was determined at random for each trial, and remained constant throughout a trial. The simulated distance of the ellipsoid on the right could be set to any value between 40 and 200 cm. Subjects could independently vary the width and depth of either ellipsoid, the depth being the thickness along the line of sight, and the width being its angular subtense (the height was always identical to the width).

In some cases the ellipsoid on the left rotated sinusoidally up and down around a horizontal axis (0.25 Hz; $\pm 15°$). 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 table-top and wall were covered with black cloth to reduce reflection, and the stimuli were red and quite dim (about 0.9 cd/m^2 as seen through the spectacles and red filters). An additional advantage of using red stim-

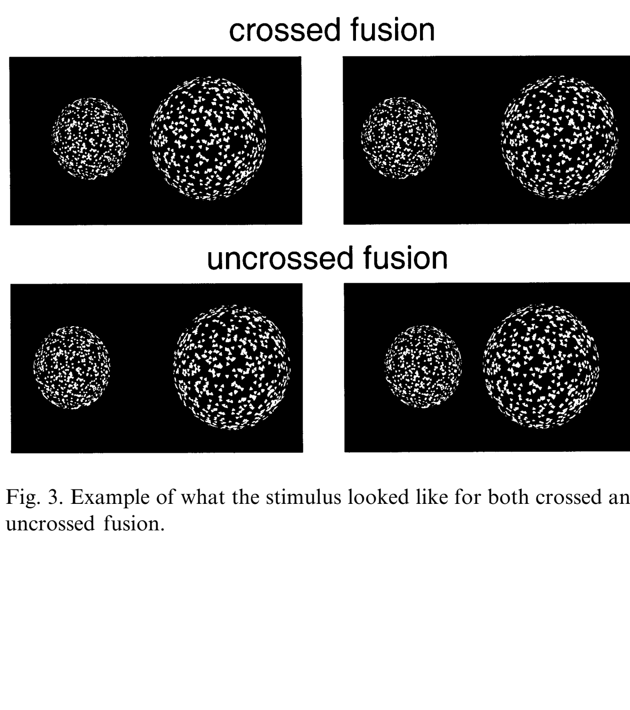

Fig. 3. Example of what the stimulus looked like for both crossed and uncrossed fusion.

uli is that red light hardly stimulates the rods, which reduces the effective change in sensitivity during dark adaptation.

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 inconsistency in the stimulus was a conflict with accommodation (for all but the parts of the ellipsoid that were at a simulated distance of 80 cm). Shading provided no useful information (surfaces were rendered with uniform illumination).

2.3. Procedure

The subjects' task was to set the size and shape of the simulated ellipsoids to match a tennis ball (radius = 33 mm). They were also required to set the distance of the ellipsoid on the right to half that of the one on the left. During the experiments they held a real tennis ball in their left hand and the computer mouse in their right hand. Subjects were encouraged to look at the tennis ball before each session, but they were not allowed to do so during the session. They adjusted the simulated ellipsoid's width and depth by moving the computer mouse. Horizontal mouse-movements changed the width of the simulated ellipsoid. The radius could vary between 0.5 and 6°. Vertical mouse-movements changed its depth. Disparity between virtual points at the ellipsoid's centre and at its nearest position could vary between 0.01 and 1.5°; we refer to the axis that is along the line of sight in the static conditions as the depth of the ellipsoid.

The simulated distance of the left ellipsoid, the initial width and depth of each ellipsoid, and the initial distance of the right ellipsoid were all determined at random (from within their entire ranges) for each trial. Subjects could choose which ellipsoid they were setting by pressing the corresponding button of the computer mouse. They could choose to set the distance of the right ellipsoid by pressing the central mouse button. In that case vertical mouse-movements brought the ellipsoid nearer or further away from them (along the line of sight; changing d_2 in Fig. 2). The simulated dimensions (in cm) were not modified when doing so, unless this would result in values going beyond the ranges given above. This did occasionally happen when subjects decreased the distance of the right ellipsoid before setting its size[2]. Subjects could switch between the different attributes that they were supposed to set as often as they liked. They indicated that they were satisfied with the settings by pressing the space bar.

In the first additional experiment, subjects were again asked to set both ellipsoids to match a tennis ball, but they were now asked to set the right ellipsoid to the same distance as that on the left. The range of distances of the ellipsoid on the left was not changed, but the range of distances that could be set for the ellipsoid on the right was changed to 60–240 cm to accommodate the larger distances that subjects were expected to set. There was only one condition: the ellipsoid on the left was always rotating.

The second additional experiment was identical to the first, but the subjects were asked to set the size of the ellipsoid on the right to twice the size of a tennis ball. This allowed us to present the two simulated objects at the same distance, while making it more difficult to compare the retinal disparities directly. The purpose of these additional experiments was to replicate Economouly and Landy's (1995) findings with the present stimuli and procedure.

2.4. Subjects

Four subjects took part in the main experiment, including one of the authors (EB). The other three were naive as to the purpose of the experiment. All had normal binocular vision. In the main experiment, each subject made settings for between 110 and 141 trials for each of the two conditions: rotating or static left ellipsoid. They did so during two sessions of about 1 h each. Within each session, the condition was determined at random for each trial.

One naive subject's size settings were similar to those of the others, but the depths and distances she set bore no relation at all to the simulated distances. Despite having normal stereo-acuity thresholds, she did not appear to rely on stereopsis when making her settings. She often set the ellipsoid to its extreme values, although she never complained of being unable to make the setting (see van Damme & Brenner, 1997 for a report of similar behaviour in other subjects). Her data are not included in the further analysis, and she was not asked to take part in the additional experiments. The other three subjects each made 50 settings in each of the additional experiments.

2.5. Analysis

The first stage in our analysis was to plot the shapes that were set (simulated width versus simulated depth) together with curves representing the errors that would arise if binocular disparity and retinal extent are used to judge the shape, but the distance used to interpret them is incorrect. These curves are based on the ap-

[2] In the main experiment, subjects were explicitly told that they were not supposed to set the two ellipsoids to the same distance for making the size and shape settings.

proximately quadratic relationship between width and depth as a function of distance (see Appendix A). The curves require a measure of the size of the sphere that would be considered correct. We previously used the real size of a tennis ball as this measure (van Damme & Brenner, 1997; Brenner & van Damme, 1999). However, we detected systematic deviations (in previous as well as the present data) that suggest that individual subjects had slightly different sizes in mind. We therefore estimated the size that each subject had in mind —which we will henceforth call the subject's reference size—from their settings in the condition in which both ellipsoids were static.

2.5.1. The reference size

Assuming that depth was derived exclusively from binocular disparities, any combination of angular extent and relative disparity is consistent with a sphere of some particular size at some particular distance. The reason for this is similar to the reason for motion parallax and stereopsis together specifying the distance: the relationship between angular extent and physical size (width) is linear with distance, whereas that between retinal disparities and physical size (depth) is approximately quadratic with distance (compare Fig. 1A and C). Thus, there is a single distance at which the set angular extent and relative disparity correspond with a width and depth that are equal, as they should be for a sphere, and knowing this distance also specifies the size. Such sizes were calculated for each ellipsoid in the condition in which both ellipsoids were static, and the average of all these values was taken to be the subject's reference size.

2.5.2. Simulated and scaling distance

In order to avoid the approximations of Appendix A when conducting a more quantitative analysis, we also compared the set relative disparity (between the centre of the ellipsoid and the point nearest to the observer) with the relative disparity that would be required for the ellipsoid to become a sphere at the simulated distance, and with the relative disparity that would be required for the ellipsoid to become a sphere at the distance at which the angular extent that the subject set would match the reference size.

This gives us two predictions for the relative disparity that subjects set: one in which the perceived distance is fixed and veridical (*simulated distance*) and the size emerges from combining this distance with the angular extent; and a second in which the size is fixed (reference size) and the distance emerges from combining this size with the angular extent (*scaling distance*). We calculated the difference between the predicted and set disparity with respect to each of these measures of distance, and used this to determine which prediction was better for each setting.

2.5.3. Statistical evaluation

Most of the effects were so clear that no statistical evaluation was deemed necessary. However, we did analyse the frequencies with which each measure of distance (*simulated* or *scaling* distance) resulted in better predictions. The distribution in each of the three conditions in which one of the ellipsoids was rotating was compared with that when both were static (for each ellipsoid) using the χ^2 test for two independent samples (Siegel, 1956). Any P value smaller than 5% was considered significant.

2.5.4. Set distance

We also analysed the distance at which the right ellipsoid was set. To do so we determined the relative disparity between the centres of the two ellipsoids. In analogy to the analysis of the relative disparity within each ellipsoid, the set relative disparity between the two ellipsoids was compared with the disparity that would set the right ellipsoid to half of the simulated distance of that on the left, and with the disparity that would set the right ellipsoid to half of the scaling distance of the ellipsoid on the left.

We also determined the ratio between the simulated distances of the (centres of the) two ellipsoids on each trial. These ratios are expected to group closer to 0.5 as judgements of distance become more veridical[3]. Moreover, we computed the ratio of the angular widths of the two ellipsoids on each trial (which is approximately the same as the ratio of the scaling distances of the two objects). This ratio is always expected to be close to 0.5, because the same size at half the distance results in twice as large a retinal extent, irrespective of the distance. The two ellipsoids' order was reversed for the distance ratio (d_2/d_1 in Fig. 2) and the ratio of angular widths $((W_1/d_1)/(W_2/d_2))$ to make the comparison easier.

3. Results

The subjects had no difficulty performing the task, but their settings were far from veridical. There were large systematic biases and there was a good deal of variability in the widths, depths and distances that were set. In order to make sense of the settings we must therefore examine the regularities in the data.

[3] Misjudging the distance of the left ellipsoid results in incorrect distance settings because the relative disparity that is required to halve the distance depends on the initial distance (see Brenner & van Damme, 1998).

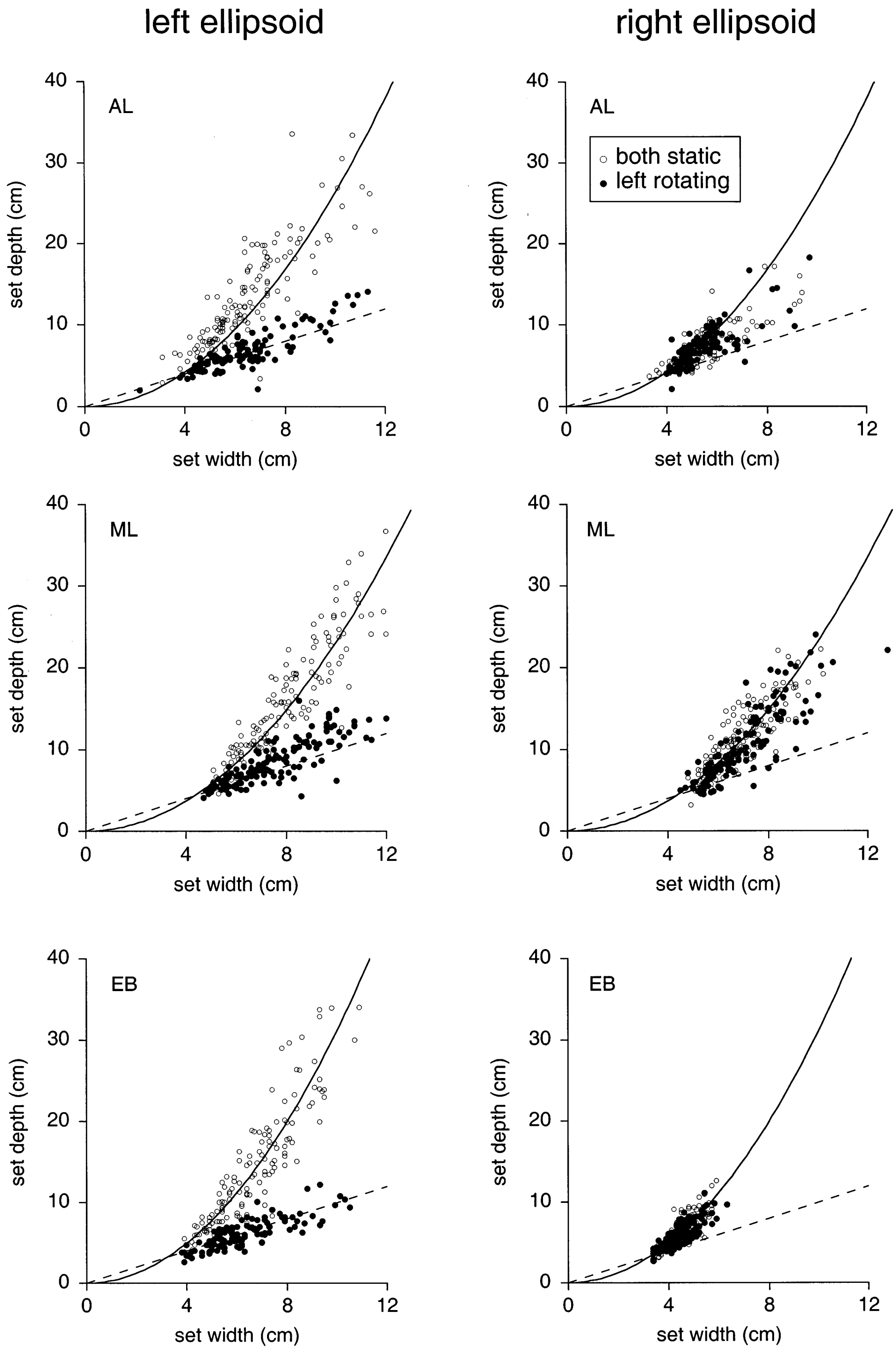

Fig. 4. Results of the main experiment for all three subjects. Set widths and depths of all ellipsoids are shown. The solid curves show the settings subjects would make if they set the disparities and retinal extent to match a sphere of a fixed radius, but misjudged the distance. The radii that were used to calculate the errors that would arise from scaling retinal extent and relative disparities with incorrect measures of distance were 3.8, 4.3 and 3.2 cm for subjects AL, ML and EB, respectively. The radius of a real tennis ball is 3.3 cm. The dashed lines represent spherical settings (i.e. a slope of 1; the set disparities and retinal extent are consistent with a sphere at the simulated distance, but the size is not necessarily correct for a tennis ball). The left panels are for the left, fixed ellipsoid, and the right panels are for the right, adjustable ellipsoid. The latter was set to half the perceived distance of the former.

3.1. Shape task

Fig. 4 shows all three subject's settings in the main experiment. Each point represents the settings for one ellipsoid. For each subject, the left panel shows the settings for the ellipsoid on the left, and the right panel shows settings for the ellipsoid on the right. Open symbols are for the condition in which both ellipsoids were static. Solid symbols are for the condition in which the left ellipsoid was rotating. The shown set

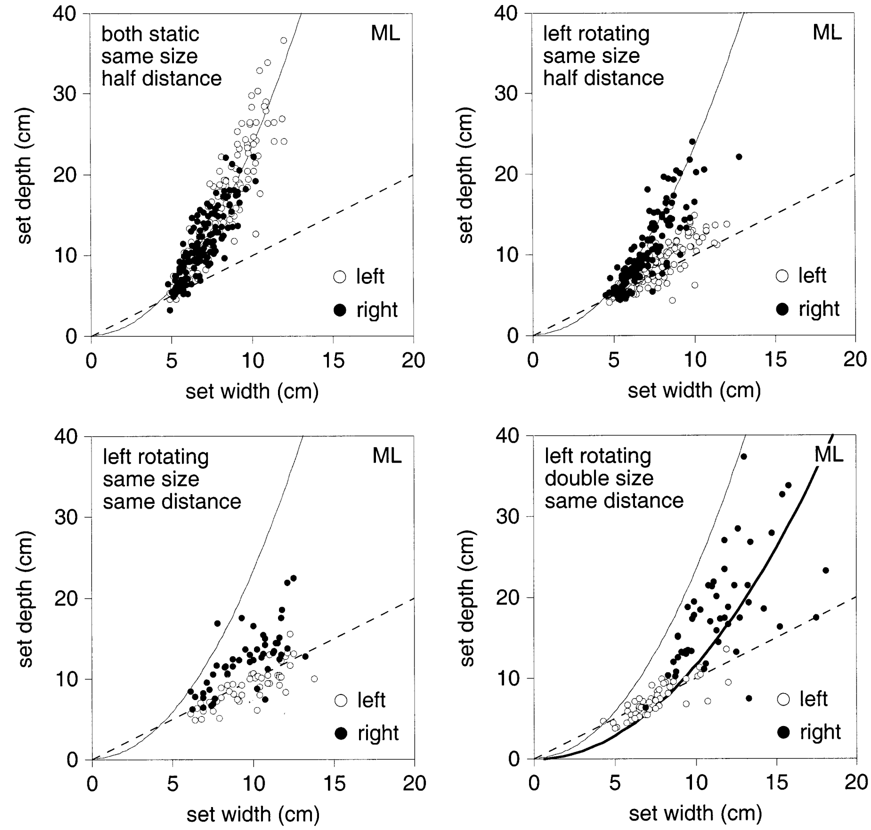

Fig. 5. One subject's data for all three experiments. Each panel shows results for one condition. The open symbols are for the left ellipsoid and the solid ones for the ellipsoid on the right. The additional thick curve in the lower right panel is the equivalent of the thin curves (in this and the preceding figure) for a twice as large ball. Other details as in Fig. 4.

values of width and depth are the simulated values, as illustrated in Fig. 2 (W_1, D_1, W_2 and D_2). The dashed line represents spherical settings (depth equal to width) at the simulated distance. If only the size were misjudged all points would lie on this line.

According to our calculations, both the naive subjects overestimated the size of a tennis ball, despite holding one in their hand. The calculated reference size was 3.8 for subject AL, 4.3 for subject ML and 3.2 for subject EB. The curves show the kind of errors one would expect (for a ball of this size) if only the distance is misjudged, with depth settings being based exclusively on binocular disparities[4]. It is evident that misjudging the distance is indeed an important source of the variability when the ellipsoid is static (open symbols).

The picture was quite different when the ellipsoid was rotating (solid symbols in the left panels). As was to be expected under these conditions (Johnston et al., 1994; Econopouly & Landy, 1995; Brenner & van Damme, 1999), rotating the ellipsoid clearly influenced the set depth, making the set shape much more veridical. This improvement in set shape was not accompanied by an improvement in set width.

The right panels show the settings for the nearer, static ellipsoid. It is difficult to tell from the figures whether there is any effect of the other ellipsoid (i.e. whether there is any difference between the two kinds of symbols), but it is evident that if there is such an effect, it is very modest. The settings all appear to follow the curves.

Fig. 5 shows one subject's settings in all three experiments. The panels now each show one condition, with the symbols indicating which ellipsoid is concerned. The upper two panels show the same data as in the preceding figure, but in a manner that makes it clear that there was no difference between the ellipsoids, as far as the relationship between set depth and set width is concerned, when both were static (upper left panel), but that there was a difference when one was rotating (upper right): the rotating ellipsoid—open symbols—was set to be much more spherical.

When the task was to set the two ellipsoids to the same distance (lower left panel) this subject's settings

[4] Systematically misjudging the size of a tennis ball has already been accounted for by using the calculated reference size rather than the real size of a tennis ball. The individual subject's reference size is visible in each figure as the intersection of the curve with the dashed line.

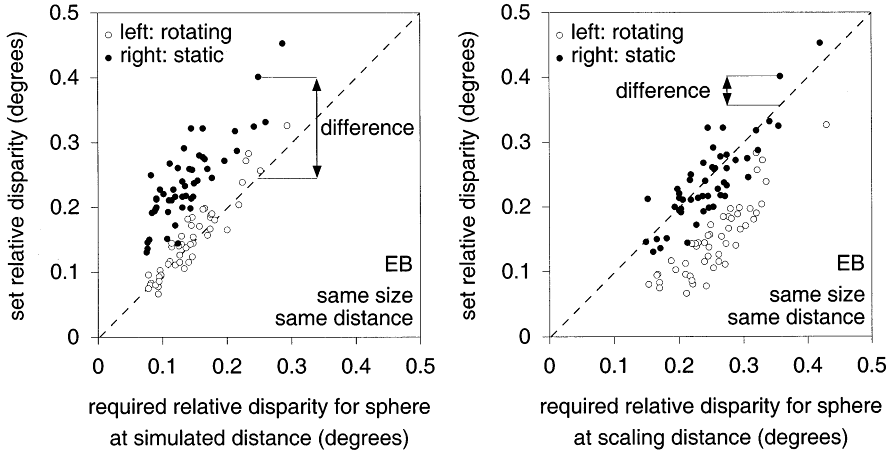

Fig. 6. The relationship between one subject's depth settings and the predictions based on two different measures of distance. All depths are expressed as the relative disparity between the centre and the nearest point of the ellipsoid. The same set depths are shown as a function of two different measures of distance. The set depth of the left, rotating ellipsoid (open symbols) is clearly more consistent with the simulated distance (left panel), whereas that of the other, static ellipsoid (solid symbols) is clearly more consistent with the scaling distance (right panel). For further analysis, the difference between the set relative disparity and each prediction was determined for each setting, and the prediction that gave the smaller difference was considered best for that setting (for the indicated setting this is the prediction based on the scaling distance). The data are for subject EB in the same size and distance task.

for the static ellipsoid (solid symbols) did deviate from the curve. The set shape was less spherical than for the rotating ellipsoid (open symbols), but there is no doubt that the rotating ellipsoid did influence the other, static one (compare the solid symbols with those in the top right panel).

When the right ellipsoid was to be set to twice the size of a tennis ball, the influence appears to be smaller (lower right panel). Note that the prediction based on misjudging distance is now slightly different (thick curve), because the size that was to be set is twice as large. It is assumed that subjects still use the same reference size (based on the static condition of the main experiment, and doubled for the ellipsoid that was to be set to twice the size of a tennis ball). The pattern of the second naive subject's settings was very similar to those shown in Fig. 5. Author EBs settings for the right ellipsoid always appeared to follow the curves.

3.1.1. Quantitative analysis

Fig. 6 shows one subject's depth settings in the additional experiment in which the ellipsoids were to be set to the same size and distance. The set depth is expressed as the relative disparity between the centre and the nearest point of the ellipsoid. The same settings are shown as a function of the predictions based on the two different measures of distance: the *simulated distance* and the *scaling distance*. For this subject it is evident that the settings for the rotating ellipsoid are more consistent with the simulated distance (which we consider the veridical distance because it is the one which introduces the least conflicts between cues; left panel), whereas the settings for the static ellipsoid are more consistent with the scaling distance (the distance at which the set angular width would correspond with the reference size; right panel). The differences between the predicted and the set disparities were used to determine which measure of distance was more consistent with each subject's settings in each condition.

The proportion of settings for which the set depth is more consistent with the simulated distance than with the scaling distance is shown for each subject and condition in Fig. 7. High values mean that the shape is set veridically for the simulated distance (i.e. the settings form a sphere), irrespective of the set size. Such values are expected, for example, if motion parallax is used to improve shape judgements directly. Low values mean that the set shape is consistent with a sphere of that subject's reference size (as far as the retinal extent and binocular disparities are concerned), but the distance may be very different from that simulated. For all three subjects, the set depths of the rotating ellipsoids were always significantly more consistent with the simulated distance than were the set depths of the static ones (open bars; $P < 0.001$). When it was the other ellipsoid that was rotating (solid bars), this was not always the case.

Subject ALs data show no influence of the rotation unless the two ellipsoids were set to the same distance and size (in which case the settings became significantly more consistent with the simulated distance; $P < 0.001$). The second naive subject's settings became significantly more consistent with the simulated distance whenever the other ellipsoid was rotating (ML; $P < 0.001$), but the change was not striking unless the two were set to the same distance and size. The third, non-naive sub-

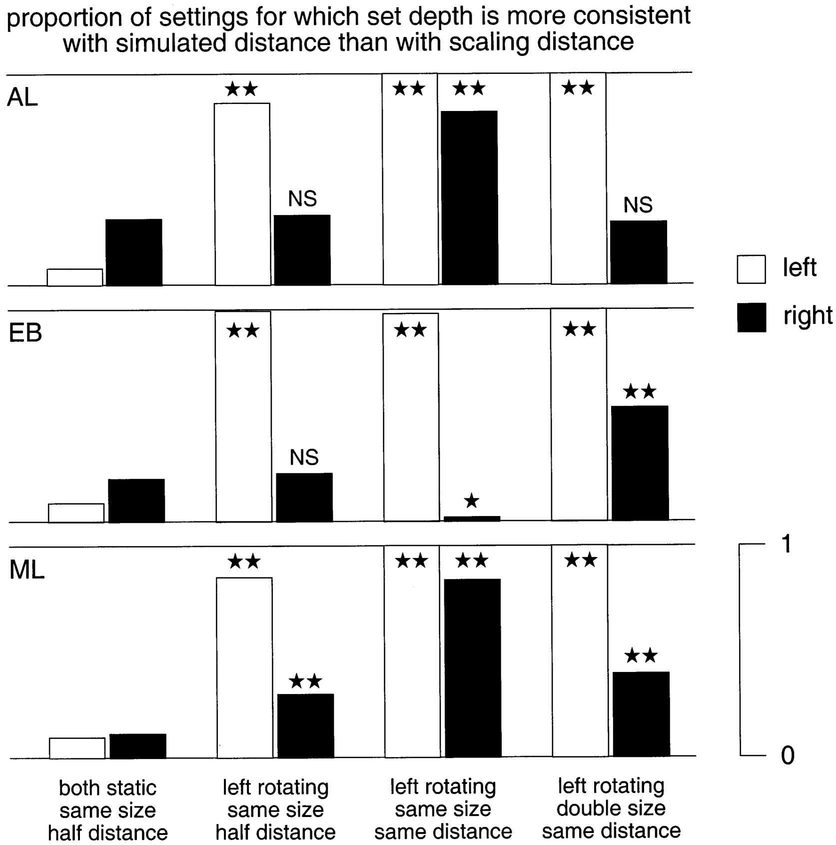

Fig. 7. Summary of a comparison between the two predictions for all four conditions and all three subjects. Open bars are for the left ellipsoid and solid ones for the right ellipsoid. Each proportion was compared with the proportion for the corresponding ellipsoid in the condition in which both ellipsoids were static using χ^2 tests for two independent samples. NS, $P > 0.05$; *, $P < 0.01$; **, $P < 0.001$.

ject's settings were not affected when the ellipsoids were at different distances. When they were at the same distance they were less consistent with the simulated distance if they were the same size[5] (EB; $P < 0.01$), but more consistent if they were of different sizes ($P < 0.001$).

3.2. Half distance task

We expected subjects to set the relative disparity between the two ellipsoids in accordance with the scaling distance of the left ellipsoid. They clearly did not. Fig. 8 shows the set relative disparity between the ellipsoids during the half distance task for one subject. There was clearly little relationship between the set disparity and the disparity that fits either the simulated or the scaling distance of the left (fixed) ellipsoid[6]. There was also no evident difference between the conditions. The data for the other subjects (not shown) were similar.

As this way of plotting the data was not very informative, we also made histograms of the ratios of the simulated distances of the two ellipsoids, as well as of the ratios of their angular widths. Fig. 9 shows these ratios for the three subjects. Open bars are for the condition in which both ellipsoids were static. Solid ones are for the condition in which the left ellipsoid was moving. Again there was no evident difference between the conditions. Most importantly, the ratio of simulated distances was not more veridical (closer to 0.5) when the left ellipsoid was rotating. This should have been so if the judgement of distance had improved (see note 2).

We expected the ratio of the angular widths to be set almost perfectly (i.e. to a value of 0.5). This was very clearly not the case. We also expected the ratio of simulated distances to be more variable due to the variability in perceived distance, and to be displaced from the correct value (0.5) in a subject-specific manner, because systematic errors in the perceived distance of the left ellipsoid could be expected to introduce systematic errors in the set distance of that on the right. The ratio of simulated distances did indeed peak less

[5] We have no explanation for this. The influence may have been reduced by the subject being aware of the conflict in his percepts, but that would not explain the opposite effect.

[6] One might worry that subjects were influenced in their distance halving settings by running into the bottom of the range of settings allowed them (40 cm). However, only five of the distance settings for the right ellipsoid (all for subject AL) were within 2 cm of the limits of the range that was provided, and most were far from these limits.

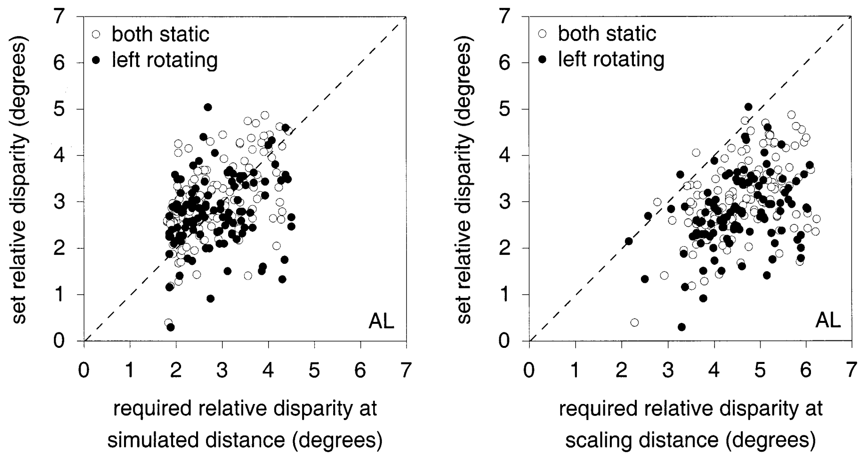

Fig. 8. The relationship between one subject's distance settings for the right ellipsoid and the predictions based on the two different measures of distance for the left ellipsoid (in the half distance task). The two conditions are indicated by the symbols. Distance is expressed as the relative disparity between the centres of the two ellipsoids. Neither prediction is very good. There is no evident difference between the conditions.

sharply than the ratio of angular widths, but the most evident systematic deviation from 0.5 was found for the ratio of the angular widths rather than of the simulated distances: this ratio was clearly larger than 0.5 for all three subjects.

4. Discussion

Despite the approximations mentioned in the first paragraph of the analysis section (and in Appendix A), the curves in Figs. 4 and 5 provide a useful indication of the kind of errors one would expect if only the distance were misjudged. It is evident from the settings that the ellipsoids' distances were indeed misjudged, and that this gave rise to systematic errors in the set shapes (for static ellipsoids).

In the present study, we used a stimulus and degree of rotation which we knew (Johnston et al., 1994; Econopouly & Landy, 1995; Brenner & van Damme, 1999) would give rise to almost veridical judgements of shape. And indeed, adding shape information to the left ellipsoid, by rotating it, resulted in the set shape of that ellipsoid becoming much more spherical. When combining information from motion parallax and binocular stereopsis to judge shape, subjects certainly do not always attribute more weight to motion parallax (see Tittle & Perotti, 1997). They do not even necessarily do so in conditions in which motion parallax clearly provides more veridical information (see experiment 1 in Tittle et al., 1995). The reason for the dramatic improvement in the perceived shape of an object when it is rotated in some studies (such as the present one, and Johnston et al., 1994), but not in others (such as Tittle et al., 1995), is not clear. However it is also not too surprising, because many factors may affect the extent to which subjects rely on each of the conflicting cues. In these specific cases the difference may have to do with the amplitude of rotation, the rate at which new images were presented, conflicts with other cues (such as texture), fixation requirements, or whether the visible contours belong to the object or to an aperture through which the object is seen[7]; all of which differed between the studies. A control experiment that is presented in Appendix B shows that local motion information was critical for the improved judgements of shape in our rotating stimuli (i.e. that the improvement was not solely due to detecting changes in the contour, or in local stereoscopic depth or orientation).

For the conclusions we want to draw from the present experiment, the reason for the improved shape judgements is irrelevant. What is important is that subjects set the shape of the rotating ellipsoid to be spherical, despite the fact that this results in binocular disparities that the subject would not accept as belonging to a sphere if the ellipsoid were not rotating. The fact that the improved shape judgements when the sphere is rotating did not give rise to better judgements of that ellipsoid's size, and did not influence the settings of the other ellipsoid, provides compelling evidence that subjects assigned less significance to the binocular disparities of the ellipsoid if it was rotating, rather than using the inconsistent disparities to improve their estimates of the viewing distance. This finding is consistent with earlier work (Brenner & van Damme, 1999). The main new finding is that the set shape of the other ellipsoid was not influenced when the two ellipsoids

[7] Motion parallax may normally be combined with other information, such as information about the position of—or changes in—the contour, so that masking the contour may make motion parallax less effective for judging shape (see Shigemasu & Sato, 1998).

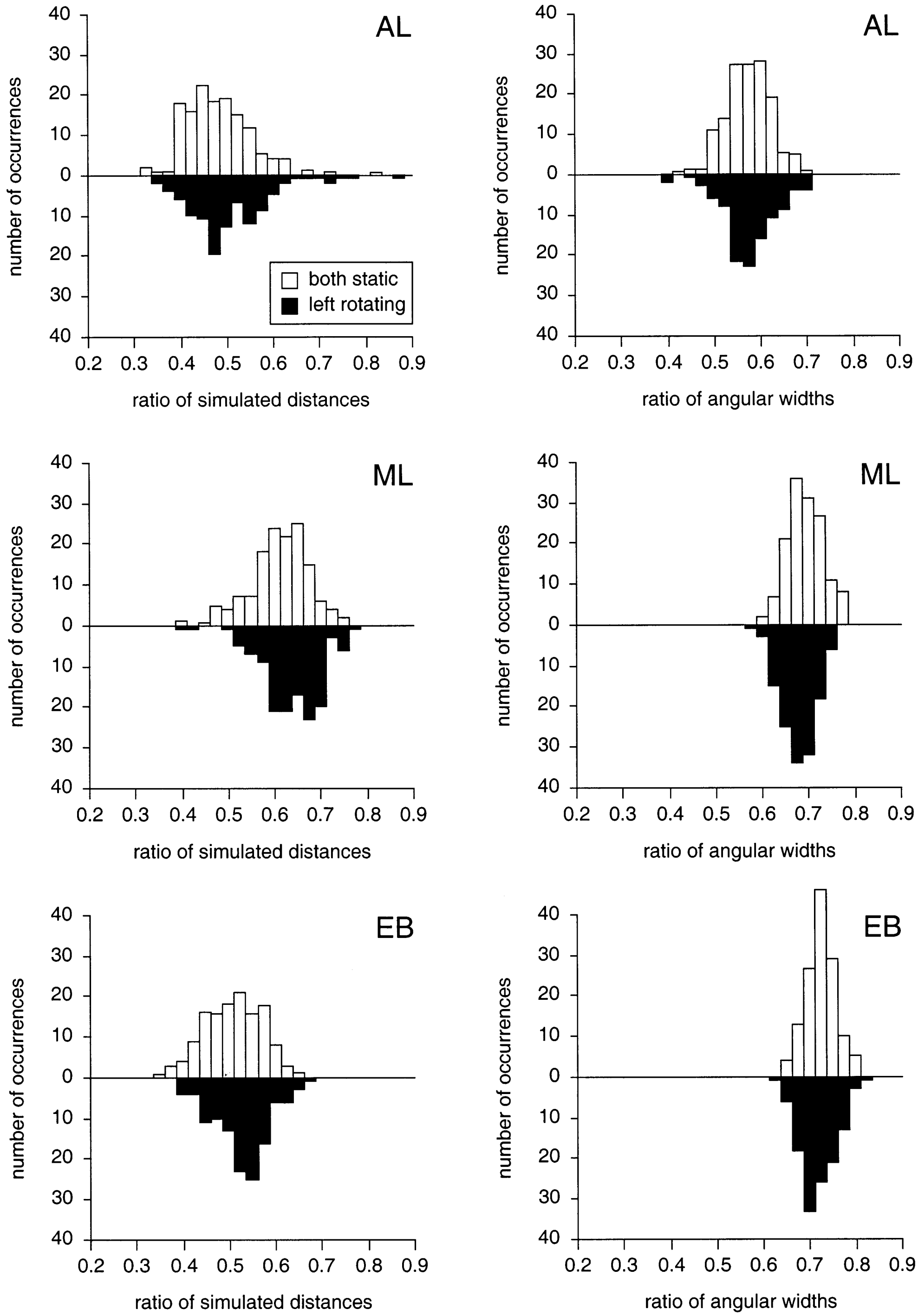

Fig. 9. Distribution of the ratios of the simulated distances and of the angular widths in the half distance task. The former shows how inaccurately subjects set the right ellipsoid to half the simulated distance of the one on the left. The latter indicates that subjects did not even set the retinal extents to differ by a factor of two. Open (upward) bars are for the condition with both ellipsoids static and solid (downward) ones are for the condition in which the left ellipsoid is rotating.

were at different distances (right panels in Fig. 4). This allows us to conclude that subjects do not combine shape cues to obtain more veridical information about distance.

A quantitative analysis of the data is far from trivial. Any measure one takes is based on certain assumptions, which may or may not be correct. We limited our analysis to choosing between two measures of distance: the *simulated distance* (consistent with vergence, texture, and subtle binocular cues such as vertical disparities and the extents of the monocular regions) and a hypothetical *scaling distance*, which is the distance that is used to relate retinal extent to actual physical dimensions (van Damme & Brenner 1997; Brenner & van Damme 1999). To be able to compare these two measures, we expressed the deviations of the set depth from the depth predicted by each measure in degrees of relative disparity. This does not eliminate all complications (see next paragraph), but it provides us with a compact way of summarising the data (Fig. 7).

Most of the complications arise because although both predictions specify the viewing distance, the prediction based on the scaling distance also specifies the size (the hypothetical reference size), whereas the prediction based on the simulated distance does not (it only specifies that the object should be a sphere). Thus, because the reference size was based on the trials in which both ellipsoids were static, one may expect a slightly better fit to the scaling distance (which depends on the reference size) in that condition. Similarly, variability in the assumed size of a tennis ball during the experiments, changes in this reference measure between experiments, or errors in setting the size for some other reason, will make the scaling distance appear less appropriate (because this distance depends on the set size), but need not influence the deviations from the prediction based on the simulated distance (because only the relationship between set width and depth is considered; the size may be quite incorrect).

Thus, variability in the reference size may make the settings appear to be more consistent with the simulated distance than they actually were. This is likely to be least so when both ellipsoids were static (because the reference size was determined from these data) and most clearly so when the ellipsoid was set to twice the perceived size of a tennis ball (because subjects may not have been aiming for twice the reference size). Note that neither of these biases contradicts our conclusions. On the contrary, they may explain some of the modest effects that are not consistent with the conclusion.

The results of the experiments in which the two ellipsoids were to be set to the same distance are consistent with the findings of Econopouly and Landy (1995). Fig. 7 can be considered to show the extent to which subjects made spherical settings for the simulated distance. For the right ellipsoid, this tendency was most evident (at least for the two naive subjects) when the two ellipsoids were at the same distance and were to be set to the same size. In that case it is not unreasonable for there to be a tendency to set the same disparity in both objects, because the disparity should be the same for two objects of the same shape and size. Some of Econopouly and Landy's findings can be interpreted as a tendency to match disparities when the targets appear to have the same shape[8]. In both studies this is only a tendency. The disparities are clearly not matched. When the task was to set the left ellipsoid to twice the size of that on the right, the tendency to make spherical settings was less clear, but did not disappear altogether. This is (at least qualitatively) consistent with Econopouly and Landy's finding that the two objects did not have to appear to be the same size for an influence to be found.

The influence of the rotating ellipsoid on the second, static one was most evident when the two ellipsoids were to be set to the same size and distance. It was smaller when the right ellipsoid was to be set to twice the size, and (almost) disappeared when the right ellipsoid was to be set to half the distance. It would appear therefore that subjects were being influenced by a direct comparison of the disparities. If so, subjects must have been comparing the set disparities of the rotating ellipsoid with those of the other ellipsoid (giving rise to the more spherical settings for the static ellipsoid), despite ignoring these same disparities when setting the rotating ellipsoid's shape (the set disparities for the rotating ellipsoid were often clearly outside of the range that the subjects would consider acceptable if the ellipsoid had not been rotating).

The half-distance settings are difficult to interpret. We know that subjects can make quite reliable half-distance settings for much simpler stimuli (Brenner & van Damme 1998). The set disparity in the present study was neither consistent with the simulated nor with the scaling distance. Even more surprisingly, the ratio of the set angular widths was much larger than the factor of two that one would expect. This implies that the perceived distance is not equal to the scaling distance, because the perceived distance was presumably halved (in accordance with the task) while the scaling distance evidently was not (because that would result in a ratio of 0.5 for the angular widths).

It is unlikely that the misperceived distances arise from the conflict with accommodation (or any other factor that would reduce the range of perceived distances). Subjects' tendency to underestimate the simulated distance, and therefore to set too large widths and depths, could be considered to be consistent with an influence of accommodation on the perceived distance of the left ellipsoid. The left ellipsoid was simulated at between 80 and 200 cm, while accommodation always indicated a distance of 80 cm. The average scaling distance was 76 cm (AL: 82; ML: 77; EB: 70). However, accommodation also indicates

[8] They used cue conflict stimuli with different depths specified by relative disparity and motion parallax. Several combinations that gave rise to the same perceived shape were compared. There was a clear tendency to make a second, static object flatter if little of the perceived depth of the rotating one was due to relative disparity.

a distance of 80 cm for the right ellipsoid. If this had influenced the perceived distance appreciably, then in order to make the right ellipsoid appear to be at half the distance, the subjects would have to overcome a tendency of accommodation to make it appear further away, so that the scaling distance for the right ellipsoid would have to be less than half of that for the left ellipsoid (less than 38 cm). In fact, the average scaling distance was 50 cm (AL: 47; ML: 53; EB: 50).

It is well-known that angular extent influences the perceived distance of isolated objects in the dark, even if one is not familiar with the object's size (e.g. Brenner, van den Berg & van Damme, 1996; Brenner, van Damme & Smeets, 1997). A possible explanation for the peculiar pattern of set distances in the present study is therefore that the angular extent influenced the perceived distance directly, without affecting the scaling distance. This could explain how the ratio of scaling distances (and therefore of the angular widths) can be larger than 0.5, while the ratio of perceived distances is 0.5. The nearer ellipsoid looks closer than the combination of relative disparity and scaling distance would suggest—due to the larger retinal image—whereas the more distant ellipsoid looks further away—due to the smaller retinal image. This raises the paradoxical conclusion of there being a measure of distance that is considered suitable to scale retinal extent and disparities (van Damme & Brenner, 1997; Brenner & van Damme, 1999), but that is not considered good enough to fully determine the perceived distance.

Acknowledgements

We thank Larry Maloney and Ivan Lamouret for their helpful comments. Michael Landy was supported in part by grant EY08266 from the National Institute of Health.

Appendix A

The curves in Figs. 4 and 5 are based on several approximations. The difference in the direction to any structure from the two eyes (α; see Fig. 10) is:

$$\alpha = \arctan\left(\frac{d \times \cos(\beta) + h}{d \times \sin(\beta)}\right) - \arctan\left(\frac{d \times \cos(\beta) - h}{d \times \sin(\beta)}\right)$$

where d is the distance to the structure (from a point midway between the eyes), h is half of the subject's inter-ocular distance, and β is the structure's horizontal direction. This can be written as:

$$\alpha = \arctan\left(\frac{2hd \times \sin(\beta)}{d^2 - h^2}\right)$$

which, at much larger distances than the inter-ocular distance (h^2 negligible in comparison with d^2), is approximately:

$$\alpha = \arctan\left(\frac{2h \times \sin(\beta)}{d}\right).$$

The set relative disparity (δ) between the centre and the nearest point on the ellipsoid is therefore approximately:

$$\delta = \arctan\left(\frac{2h \times \sin(\beta)}{d}\right) - \arctan\left(\frac{2h \times \sin(\beta)}{d - D}\right)$$

where D is the set depth.

Assuming that subjects interpreted this relative disparity as arising from a ball with radius T at a distance d', and that they do not misjudge retinal disparities,

$$\arctan\left(\frac{2h \times \sin(\beta)}{d}\right) - \arctan\left(\frac{2h \times \sin(\beta)}{d - D}\right)$$
$$= \arctan\left(\frac{2h \times \sin(\beta)}{d'}\right) - \arctan\left(\frac{2h \times \sin(\beta)}{d' - T}\right)$$

However, if subjects interpret the image to be at distance d', then the set width (W) should also be related to the size of the ball (T) by the relationship:

$$\frac{W}{d} = \frac{T}{d'}.$$

Combining the two preceding equations to get rid of d', and solving for D using a similar approximation to that used above [$(2h \times \sin(\beta))^2$ must now be negligible in comparison with $d \times (d - D)$ and $d' \times (d' - T)$], we get:

$$D = \frac{dW^2}{Td - TW + W^2}$$

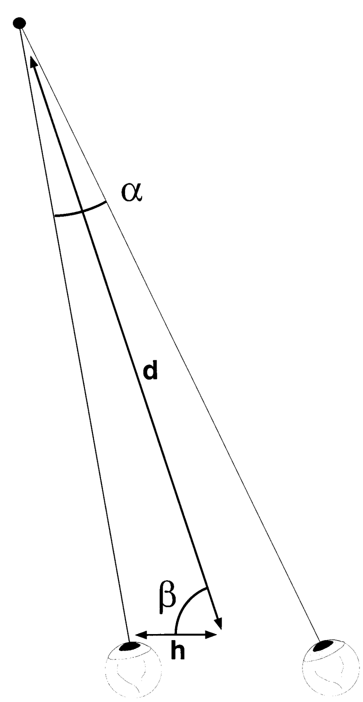

Fig. 10. Schematic representation of several angles and distances that can be used to describe the relationships between the two eyes and an arbitrary point in space.

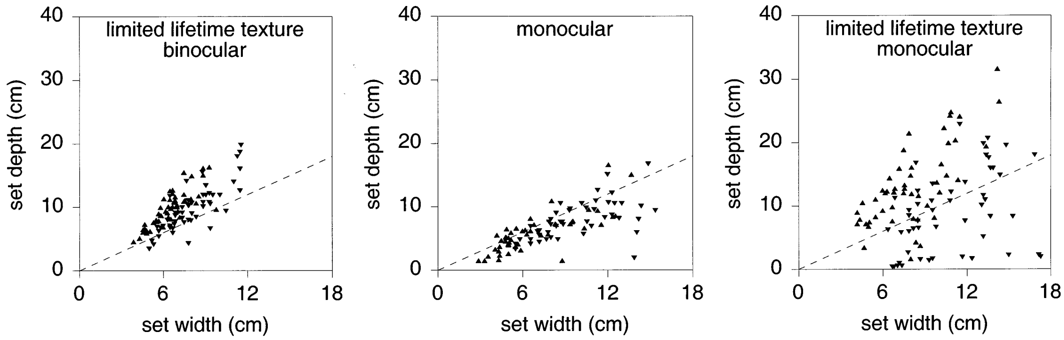

Fig. 11. Results of the monocular and limited lifetime experiment. Set widths and depths of all ellipsoids are shown. The dashed lines represent spherical settings. Each panel shows a different condition. The two subjects are identified by symbols: ▲ EB; ▼ ML.

which, as long as $W-T$ is not too large (i.e. as long as the error in set width is small in comparison to the simulated distance, d) is approximately

$$D=\frac{W^2}{T}$$

which is the relationship that is shown in the figures. The attractive aspect of these approximations (for which the inter-ocular distance, the set size, and the error in set size must be small in comparison with the simulated distance) is that they get rid of the simulated distance, allowing us to draw a single curve in the figures. Assuming that subjects are more likely to systematically misjudge the size of a tennis ball, than to systematically misjudge its shape, we decided not to assume that T, the perceived radius of the set ball, is 3.3 cm (corresponding to a real tennis ball). Instead, we estimated T from the data (assuming that the percept was always spherical). For individual trials h, β, d, D and W are known, so that T (and d') can be determined without having to make any of the above-mentioned approximations. Doing so for all trials in which both ellipsoids were static gives us a large number of estimates of T. The average of these values (for each subject) was used as that subject's *reference size*. This is the value of T that is used in the curves in Figs. 4 and 5, and is the value that determines the *scaling distance*. This scaling distance is the value of d' from the relationship

$$\frac{W}{d}=\frac{T}{d'}$$

when T is the reference size, d the simulated distance, and W the set width.

Appendix B

Beside introducing motion parallax as a depth cue, rotating the ellipsoid also gave rise to changes in contour when the ellipsoid was not spherical. One might therefore attribute the improved shape judgements to a simple strategy of keeping the occluding contour shape static (and circular). Rotating the ellipsoid also gives rise to local changes in depth and orientation. When the near end of an elongated ellipsoid is pointing downward, its surface at any distance below the centre is nearer than that at the same distance above the centre, so that the surface at the centre is slanted upwards. Obviously, when the near end is pointing upward, the converse is true. Thus, as the ellipsoid rotates, the depths and orientations at each position within the image change.

The existence of such cues does not affect our reasoning, because even if they were the sole cause of the improved shape judgements, reconciling the set disparities with the spherical shape would require a change in the assumed viewing distance. Nevertheless, we conducted an additional experiment to evaluate the contribution of motion parallax to the improvement.

Two of the same subjects made 150 settings, split over three conditions, for a single rotating target straight in front of them. The simulated distance was varied as in the main experiment. The task was to match the tennis ball. A *monocular* condition consisted of showing the usual image to the right eye, but nothing to the left eye. Two *limited lifetime texture* conditions were identical either to the rotating ellipsoid in the other experiments or to the monocular condition of this experiment, but the texture on the ellipsoid was replaced every 50 ms (keeping the texture lifetime below the threshold for detecting structure from motion; Treue, Husain & Andersen, 1991).

The settings are shown in Fig. 11. The settings for the *limited lifetime texture-binocular* condition are more spherical than for a static ellipsoid (compare with Fig. 4; open symbols), showing that changing contour, changing local depth, changing local orientation, or some combination of these cues can help observers detect that an ellipsoid is not spherical. However, the settings are less spherical than for a normal rotating ellipsoid (Fig. 4; solid symbols), showing that motion parallax (the only information that is missing in this condition) also helps determine whether the ellipsoid is spherical. In this condition the changing disparities often made the ellipsoid appear to undergo systematic deformations, rather than to rotate. In the *monocular* condition, in which subjects did have motion parallax information, the settings were comparable to those for binocular rotating ellipsoids. In the *limited lifetime texture-monocular* condition, in which subjects had the same information from texture and from the changing contour, but no information from motion parallax, the stimulus never looked spherical. The subjects could make settings on the basis of the contour, but they were clearly aware of doing so (the image always looked flat), and the settings were much poorer than those in the monocular condition. Thus, the improvement in the judged shape of rotating ellipsoids in the main experiments was probably primarily the result of information from motion parallax.

References

Brenner, E., & van Damme, W. J. M. (1998). Judging distance from ocular convergence. *Vision Research*, *38*, 493–498.

Brenner, E., & van Damme, W. J. M. (1999). Perceived distance, shape and size. *Vision Research*, *39*, 975–986.

Brenner, E., van den Berg, A. V., & van Damme, W. J. (1996). Perceived motion in depth. *Vision Research*, *36*, 699–706.

Brenner, E., van Damme, W. J. M., & Smeets, J. B. J. (1997). Holding an object one is looking at: kinesthetic information on the object's distance does not improve visual judgements of its size. *Perception and Psychophysics*, *59*, 1153–1159.

van Damme, W., & Brenner, E. (1997). The distance used for scaling disparities is the same as the one used for scaling retinal size. *Vision Research*, *37*, 757–764.

Econopouly, J. C., & Landy, M. S. (1995). Stereo and motion combined rescale stereo. *Investigative Ophthalmology and Visual Science*, *36*, S665.

Foley, J. M. (1980). Binocular distance perception. *Psychological Review*, *87*, 411–434.

Foley, J. M. (1985). Binocular distance perception: Egocentric distance tasks. *Journal of Experimental Psychology*: *Human Perception and Performance*, *11*, 133–149.

Frisby, J. P., Buckley, D., Wishart, K. A., Porrill, J., Gårding, J., & Mayhew, J. E. W. (1995). Interaction of stereo and texture cues in the perception of three-dimensional steps. *Vision Research*, *35*, 1463–1472.

Gillam, B. J. (1968). Perception of slant when perspective and stereopsis conflict: experiments with aniseikonic lenses. *Journal of Experimental Psychology*, *78*, 299–305.

Glennerster, A., Rogers, B. J., & Bradshaw, M. F. (1996). Stereoscopic depth constancy depends on the subject's task. *Vision Research*, *36*, 3441–3456.

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., Maloney, L. T., Johnston, E. B., & Young, M. (1995). Measurement and modeling of depth cue combination: in defense of weak fusion. *Vision Research*, *35*, 389–412.

Livingstone, M., & Hubel, D. (1988). Segregation of form, color, movement, and depth: anatomy, physiology, and perception. *Science*, *240*, 740–749.

Marr, D. (1982). *Vision*. San Francisco, CA: W.H. Freeman and Company.

Richards, W. (1985). Structure from stereo and motion. *Journal of the Optical Society of America A*, *2*, 343–349.

Rogers, B. J., & Collett, T. S. (1989). The appearance of surfaces specified by motion parallax and binocular disparity. *The Quarterly Journal of Experimental Psychology*, *41A*, 697–717.

Shigemasu, H., & Sato, T. (1998). Depth cue interaction between binocular stereopsis and structure-from-motion. *Investigative Ophthalmology and Visual Science*, *39*, S622.

Siegel, S. (1956). *Nonparametric statistics for the behavioral sciences*. New York: McGraw–Hill.

Tittle, J. S., & Braunstein, M. L. (1993). Recovery of 3-D shape from binocular disparity and structure from motion. *Perception and Psychophysics*, *54*, 157–169.

Tittle, J. S., & Perotti, V. J. (1997). The perception of shape and curvedness from binocular stereopsis and structure from motion. *Perception and Psychophysics*, *59*, 1167–1179.

Tittle, J. S., Todd, J. T., Perotti, V. J., & Norman, J. F. (1995). The systematic distortion of perceived 3-D structure from motion and binocular stereopsis. *Journal of Experimental Psychology*: *Human Perception and Performance*, *21*, 663–678.

Treue, S., Husain, M., & Andersen, R. A. (1991). Human perception of structure from motion. *Vision Research*, *31*, 59–75.

Young, M., Landy, M. S., & Maloney, L. T. (1993). A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Research*, *33*, 2685–2696.

Zeki, S., & Shipp, S. (1988). The functional logic of cortical connections. *Nature*, *335*, 311–317.