PERGAMON

Vision Research 39 (1999) 975–986

Vision Research

Perceived distance, shape and size

Eli Brenner *, Wim J.M. van Damme

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

Received 28 April 1997; accepted 3 October 1997

Abstract

If distance, shape and size are judged independently from the retinal and extra-retinal information at hand, different kinds of information can be expected to dominate each judgement, so that errors in one judgement need not be consistent with errors in other judgements. In order to evaluate how independent these three judgements are, we examined how adding information that improves one judgement influences the others. Subjects adjusted the size and the global shape of a computer-simulated ellipsoid to match a tennis ball. They then indicated manually where they judged the simulated ball to be. Adding information about distance improved the three judgements in a consistent manner, demonstrating that a considerable part of the errors in all three judgements were due to misestimating the distance. Adding information about shape that is independent of distance improved subjects' judgements of shape, but did not influence the set size or the manually indicated distance. Thus, subjects ignored conflicts between the cues when judging the shape, rather than using the conflicts to improve their estimate of the ellipsoid's distance. We conclude that the judgements are quite independent, in the sense that no attempt is made to attain consistency, but that they do rely on some common measures, such as that of distance.

Keywords: Depth; Stereopsis; Motion parallax; Spatial vision; Cue interaction

1. Introduction

An intriguing question that keeps emerging in the perceptual literature is whether the way we see different aspects of our surroundings is confined to being mutually consistent. This has probably been most extensively studied for perceived size and distance (see McCready, 1985; Sedgwick, 1986; Gogel, 1990 for recent reviews), usually leading to the conclusion that the relationship between angular extent and perceived size need not be consistent with the perceived distance. Similarly, it has been shown that perceived movement is not necessarily consistent with the perceived change in position (Biguer, Donaldson, Hein & Jeannerod, 1988; Brenner, van den Berg & van Damme, 1996). Such inconsistencies appear to support the notion of independent modular processing of visual information (Marr, 1982; Bruno & Cutting 1988), especially considering that different aspects of the visual information appear to be used independently for guiding actions (Jeannerod 1981; Brenner & Smeets, 1996a,b; Paulignan & Jeannerod, 1996).

Striving for mutual consistency can be seen as an attempt to find a single physical interpretation for simultaneous judgements about diverse aspects of our surrounding (Rogers & Collett, 1989). In many of the previous studies, the judgements of different aspects were not simultaneous. Subjects are known to evaluate the same scene differently when doing different tasks (Arend & Reeves, 1986; Cornelissen & Brenner, 1995; van den Berg, 1996), and some misjudgements even only become evident in some tasks (Glennerster, Rogers & Bradshaw, 1996). Thus, inconsistencies between the way subjects use the information within the scene during consecutive tasks may inadvertently have been attributed to simultaneously perceiving different aspects inconsistently. Furthermore, additional unverified assumptions are usually made when deciding whether two perceived aspects are consistent. For example, much of the evidence for inconsistencies between perceived size and distance is based on the assumption that the retinal extent is registered accurately (McCready, 1985). Similarly, the discrepancies between perceived position and velocity are based on the assumption that we register time correctly.

* Corresponding author. Tel.: +31 10 4087569; fax: +31 10 4367594; e-mail: brenner@fys1.fgg.eur.nl.

0042-6989/98/$ - see front matter
PII: S0042-6989(98)00162-X

We recently devised a method to obtain simultaneous judgements of size and shape (van Damme & Brenner, 1997). Subjects' settings showed that the measure of distance that was used for interpreting the extent of the retinal image as object width was consistent with that used for interpreting retinal disparities as object depth, even though the distance itself was misjudged considerably (also see Rogers & Bradshaw, 1995). Although this could be a result of striving for mutually consistent percepts, it need not be, because if the same measure of distance is used for both judgements, and this measure is an important source of errors, the two perceptual judgements will vary consistently without this having to be explicitly imposed. It is well established that people make large systematic errors when judging distances under certain conditions (including those typically used in controlled laboratory experiments), and that other perceptual misjudgements can be related to these errors. This has been most thoroughly demonstrated for relative egocentric distances (Foley, 1980, 1985), but also appears to hold for perceived shape (Johnston 1991; Cumming, Johnston & Parker, 1991), size (Collett, Schwarz & Sobel, 1991), depth (Collett, Schwarz & Sobel, 1991) and motion (Gogel & Tietz, 1973).

Our previous study was designed to force subjects to use an estimate of distance for judging both the shape and the size. In the present study we examine how information about global shape that does not require an estimate of distance, influences the perceived size and distance. The main question was whether the perceived size and distance would also improve (in the sense of becoming more like the simulated values) because only adjustments to the estimate of distance can reconcile the binocular disparities with the improved judgement of shape. We used motion and shading to improve the judgements of shape because these cues contain information about local surface orientation that is independent of the viewing distance (assuming that contributions from the additional information provided by motion disparity (Beverley & Regan, 1973) and disparate shading (Bülthoff & Mallot, 1988) are negligible). Moreover, Johnston, Cumming & Landy (1994) have shown that motion can improve binocular judgements of shape considerably. They even found conditions in which the perceived shape when both stereopsis and motion cues were present was not a compromise between the perceived shapes when each cue was present on its own. They interpreted this in terms of improving the measure of distance with which disparities are scaled.

2. Methods

Subjects set the size and shape of a binocular simulation of a randomly textured ellipsoid to match a tennis ball. After each setting, the ellipsoid disappeared and the subject indicated where the simulation had been by holding a real tennis ball at that position (in total darkness). Our main interest was in an analysis of the consistency between size, shape and distance.

2.1. Equipment

A schematic representation of the equipment is shown in Fig. 1. The images were presented with a Silicon Graphics Onyx RealityEngine on a high resolution monitor (120 Hz; 38.4 × 29 cm; 1280 × 492 pixels; spatial resolution further 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 retinas were appropriate for an ellipsoid at the simulated distance (see Fig. 2A). Red stimuli (and additional red filters in front of the spectacles) were used because the shutter spectacles have the least cross-talk at long wave-

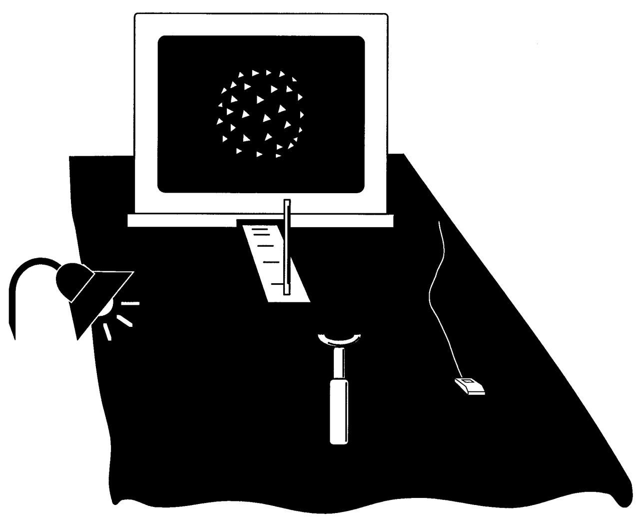

Fig. 1. Schematic view of the set-up. Subjects sat with their chin on a chin-rest, 80 cm from the monitor on which the stimulus was displayed. They changed the size and shape of the stimulus by moving the computer mouse. They indicated where they saw the stimulus by holding a tennis ball at that position (in complete darkness). The experimenter used a rod attached to a sliding ruler to measure the distance at which the subject held the tennis-ball. On some trials a table lamp provided a dim illumination of the surrounding. The table beneath the monitor and chin-rest was covered with a black cloth to minimise reflections.

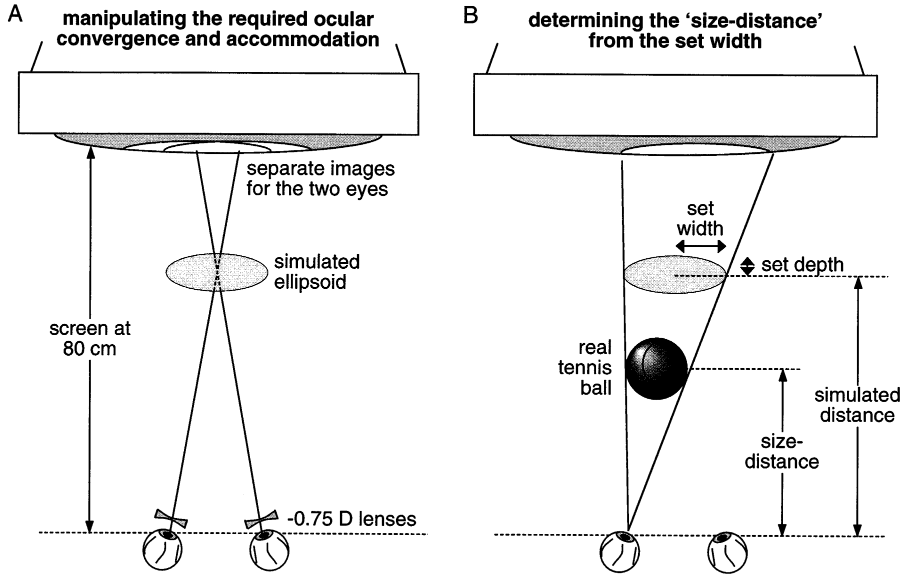

Fig. 2. Schematic view of the set-up as seen from above. (A) Alternate images were presented to the two eyes with the aid of liquid crystal shutter spectacles, so that both the ocular convergence required to fixate the ellipsoid and the binocular disparities of the triangles on the ellipsoid's surface would be adequate for an ellipsoid at the simulated position. Additional lenses were placed in front of the eyes so that accommodation would suggest a distance of 50 cm (the average simulated distance), rather than 80 cm (the distance to the screen). (B) Illustration of the way in which 'size-distance' is derived from the set width, the simulated distance, and the width of a real tennis ball.

lengths. The simulated positions of the stimuli were all in front of the screen. We placed -0.75 D lenses in front of each eye so that the required accommodation would be appropriate for the average distance that was simulated (50 cm). We did not correct for the resulting (approximately 1.4%) reduction in image size. The experimenter used a vertical rod attached to a sliding ruler to measure the distance at which the subject held the real tennis ball.

2.2. Stimuli

The stimulus was a computer-simulated opaque ellipsoid, of which only the surface texture was visible (see Fig. 3). Subjects could independently vary the size and shape of the ellipsoid. The ellipsoid's simulated distance from the observer was between 30 and 70 cm, and was determined at random from within this range for each trial. The texture on the ellipsoid's surface consisted of approximately 200 randomly oriented equilateral triangles, which were distributed at random over the surface when the ellipsoid was spherical. The triangles were all of the same simulated size, and this size did not change when the size of the ellipsoid was changed. Similarly, their shape did not change when the ellipsoid was stretched; only their positions and orientations changed. To discourage subjects from judging the size of the ellipsoid in relation to the size of the triangles, the simulated size of the latter was varied between trials. Moreover, the simulated triangle size increased in proportion to the simulated distance, so that the triangles' angular dimensions did not provide any indication of the simulated distance. An inevitable consequence is that assumptions about the simulated size may have decreased the range of perceived distances. The two variations resulted in a simulated size of the sides of the triangles of between 1.3 and 2.6 mm when the ellipsoid was at 30 cm, and between 3.0 and 6.1 mm when it was at 70 cm.

There were four conditions, with slightly different stimuli. In the first condition the subjects only saw the simulated ellipsoid. Care was taken to ensure that no other structures were visible. The table-top was covered with black cloth to reduce reflection, and the stimuli were red and quite dim. An advantage of using red stimuli is that red light hardly stimulates the rods, which reduces the effective change in sensitivity during dark adaptation. The luminance of the triangles was 0.9 cd/m^2 (as seen through the shutter spectacles and the red filter; measured with a Minolta LS-100 luminance meter). As the images were rendered in the appropriate perspective for each eye (considering individual differences in the inter-ocular distance), the stimulus contained texture cues as well as binocular disparities.

The second condition was identical to the first except that the room was dimly illuminated by a small table lamp. This was expected to improve subjects' judgements of distance because it allowed them to see other objects in the surrounding, such as the computer monitor and the surface of the table. Subjects are known to misjudge distances considerably when forced to rely on extra-retinal information about ocular convergence

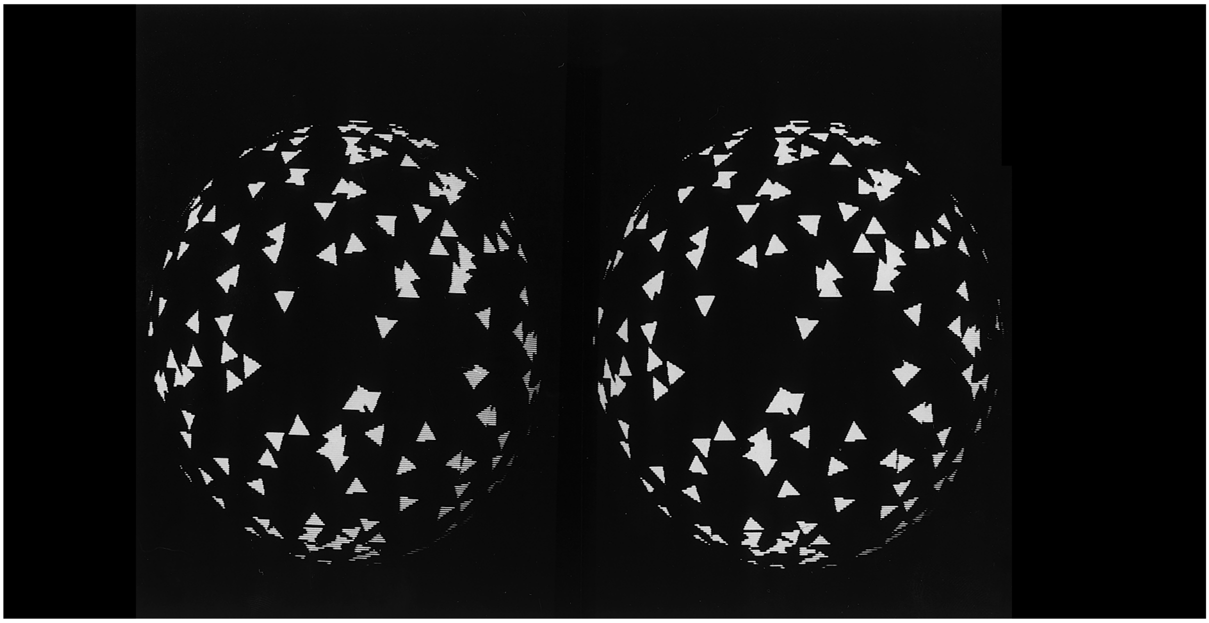

Fig. 3. Example to give an impression of what a single frame of a rotating ellipsoid looked like (for uncrossed fusion).

(Gogel, 1961; Collewijn & Erkelens, 1990; Cumming, Johnston & Parker, 1991). The distances of the familiar visible objects can presumably be determined more accurately from other sources (see Sedgwick, 1986; Cutting & Vishton, 1995), so that subjects could use the relative disparity between the ellipsoid and these objects, or even the change in vergence when shifting their gaze between the ellipsoid and these objects (Brenner & van Damme, 1998), to improve their judgements of the distance of the ellipsoid.

During the third and fourth conditions the ellipsoid rotated sinusoidally back and forth around a horizontal axis (0.15 Hz). This axis passed through the centre of the ellipsoid and was more or less orthogonal to the line of sight (i.e. the ellipsoid rotated in the sagittal plane). The ellipsoid was rendered as if the triangles were lambertian surfaces illuminated predominantly by a distant light source situated behind the subject (azimuth 45°; elevation 60°). The luminance of each triangle was determined by:

$$L = 0.34 + 0.56 \cos \alpha$$

where L is the luminance in cd/m^2 (as seen through the shutter spectacles and red filter), and α is the angle between the surface normal and the direction of the incident light. The table lamp was off.

The only difference between the third and fourth conditions was the extent of the rotation: an amplitude of either 30 or 60° from the initial direction. Johnston, Cumming & Landy (1994) showed clear effects for a rotation with an amplitude of about 15° at a viewing distance of 200 cm. Their effects were considerably less evident when they used a viewing distance of 50 cm. We were limited in our range of simulated distances by the fact that subjects had to be able to indicate them manually. We therefore increased the amplitude of the rotation, to make sure that the motion would improve the judgements of shape enough for us to decide whether the set size and indicated distance changed accordingly (Young, Landy & Maloney (1993) found that increasing the rotation amplitude increased the weight assigned to motion as a depth cue). As we were not particularly interested in isolating the motion cue, we were not worried by the additional cues that the larger amplitude may introduce. Note that shading, motion parallax, texture (in particular when a non-spherical ellipsoid is seen at an angle), and even changing disparities and a changing outline (when a clearly non-spherical ellipsoid is rotated) add information about the ellipsoid's shape without adding direct information about its distance. The smaller amplitude of rotation was included because we were worried that the ellipsoid may often look non-rigid for the large amplitude: failing to adjust the distance that is used to scale disparities could give rise to a sizeable conflict between depth cues, which, if detected, should make the ellipsoid stop looking rigid. For brevity we will often refer to the ellipsoid as rotating when referring to the last two conditions, without mentioning the shading, although the two were always present together.

2.3. Procedure

The subjects' task was to set the size and shape of the simulated ellipsoid to match a tennis ball. A tennis ball was chosen because most subjects are somewhat familiar with its size (radius = 33 mm). Subjects were also encouraged to examine the real tennis ball while receiving their instructions. During the experiments they held the tennis ball in their left hand and the computer mouse in their right hand. They were instructed to always keep the tennis ball out of sight. Subjects adjusted the depth of the simulated ellipsoid, and its angular extent, by moving the computer mouse. Horizontal mouse-movements simultaneously changed the width and height of the simulated ellipsoid (which we will refer to as its size; range: 1–7 cm), and vertical mouse-movements changed its depth (range 0.1–15 cm; we refer to the axis that is approximately along the line of sight in the static conditions as the depth of the ellipsoid). Subjects indicated that they were satisfied with their settings by pressing a mouse button. This made the ellipsoid disappear (and the lamp go off in the second condition). The Subjects then indicated where they had seen the ellipsoid by holding the real tennis ball in their left hand at that position (in total darkness). They were instructed to close their eyes when they did so, so that the experimenter could align the rod with the ball and measure the distance they indicated by the light of a small flashlight, without providing the subjects with visual information about where they held the ball. Subjects never received any feedback on their performance.

Each subject made 50 settings for each of the four conditions (during two or three sessions). Within each session, the condition was determined at random for each trial. The simulated distance, the simulated triangle size, and the initial size and depth of the simulated ellipsoid were also determined at random (from within their entire ranges) for each trial. Subjects were explicitly instructed to announce that they were unable to find a good match, or that the ellipsoid did not appear to be rigid, if this ever occurred; but no such occurrence was ever reported. Subjects received no specific instructions as to where to direct their gaze.

2.4. Subjects

Six subjects took part in the experiment. One was an author. The other five were naive as to the purpose of the experiment. All had normal binocular vision.

2.5. Analysis

The first stage in analysing the results was to try to determine the source of the misjudgements. In particular, we wanted to ascertain that errors in the judged distance were a major factor in these settings. To do so we plotted the settings together with curves representing the errors that are to be expected if one misjudges the distance. Next, we examined how the experimental manipulations influenced various measures of the errors in perceived size, shape and position. To determine whether adding information had influenced subjects' judgements, we compared individual subjects' performance in each condition in which there was additional information with their performance in the first, most limited condition (light off; static ellipsoid) with paired t-tests.

We also wanted to confirm that a single measure of distance was used for perceived size and shape, and to examine whether this coincides with the perceived distance. To do so we used a modification of the method used in our previous paper[1] (van Damme & Brenner, 1997). We first transposed the results to give us three distance values:

1. *Size-distance:* the distance at which the set retinal size would be correct for a tennis ball (see Fig. 2B).
2. *Disparity-distance*: the distance at which the range of set retinal disparities would be correct for a tennis ball.
3. *Indicated distance*: the distance indicated manually.

One may expect the three values to be the same if a single measure of distance is involved in all three judgements. This is not necessarily so, though, because each judgement is also based on certain other measures, such as measures of the retinal extent, of disparity, and of the position of the hand, each of which will introduce independent errors. Examining how improving the information about distance (by turning on the table lamp) influences the relationship between the three distance values can help determine whether the percepts of size, shape and position share a single measure of distance. If so, the relationship between the three above-mentioned values should not change. In contrast, if the information about distance is interpreted separately for each judgement, the errors need not be identical, so that the correspondence between the three values should improve when each becomes more veridical.

In order to quantify the change in the relationship between the three distance measures when the light was turned on, we determined how much was to be gained (in the sense of explaining the variability) by fitting separate lines to the relationships with the light on and off, as compared with fitting a single line to the data for both conditions together. An important feature of this analysis is that it is sensitive to differences in both the slope and the intercept of the fit lines. The two regression models were compared with F-tests for equality of the residual variances. Separate F-values were calculated for each subject and each comparison.

---

[1] Formula (3) of that paper should read

$$d_d = \frac{1}{2}R + \left(\frac{1}{4}R^2 - h^2 + \frac{R}{R_z}(d_v^2 + h^2) - d_v R\right)$$

3. Results

On average, subjects took 42 s to set the size and shape to match a tennis ball. They had no problem indicating the position of the simulation by holding the real tennis ball at that position in total darkness, but often complained about being least certain about these responses. Fig. 4 shows one subject's settings and the distances he indicated for the most limited condition (light off; static ellipsoid). The set width and height (A), and the set depth (B), were both very variable, too large (the dotted lines show the veridical values), and increased with the simulated distance. Plotting the depth as a function of the width and height (C) reveals that there is a very systematic pattern in the settings. The systematic deviations from spherical settings confirm (see Foley, 1980, 1985; Johnston, 1991; van Damme & Brenner, 1997) that a large part of the misjudgements arise from errors in judging the distance (in which case the points would follow the thick curve). This subject may also have overestimated the size of a tennis ball:

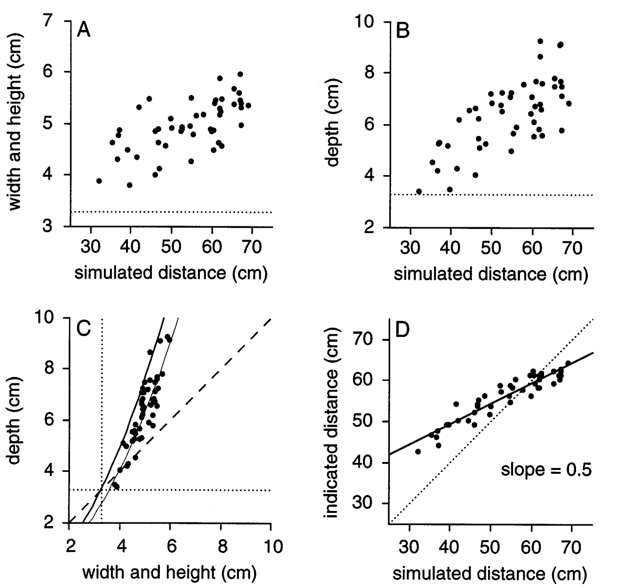

Fig. 4. Example of one subject's data in the most limited condition: static ellipsoid; light off. The width, height and depth are the set radii of the simulated ellipsoid. The indicated distance is the distance at which the subject subsequently held the tennis ball. The dotted lines indicate the width, height and depth of a tennis ball (A–C), and perfect correspondence between the simulated and indicated distances (D). The dashed, diagonal line in C indicates spherical settings. The thick curve represents the errors one could obtain by using incorrect measures of distance to interpret the retinal disparities and retinal extent: it shows approximately what subjects would set if they misjudged the distance, and set the appropriate retinal extent and range of retinal disparities for the misjudged distance. The thin curve shows the same thing for a ball with a radius of 40 rather than 33 mm. The variability in this subject's settings (A, B) appears largely to be due to his misjudging the distance (C). The solid line in D is a least squares fit to the points. It is evident that the indicated distance deviates systematically from that simulated.

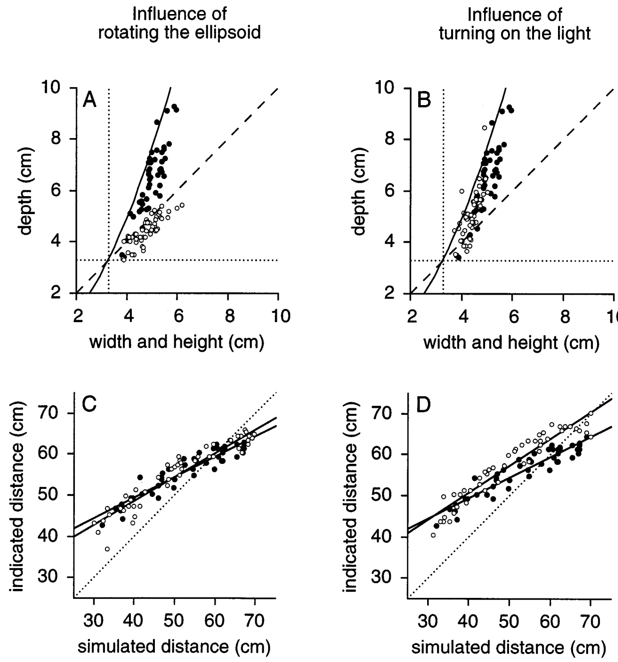

Fig. 5. Example of how adding information about shape or distance influences the judgements. The filled symbols are the same data as in Fig. 4C and D. The open symbols show the same subject's settings for the 60° rotation (A, C) and the light on (B, D) conditions. The settings for the rotating ellipsoid were more or less spherical (A). Rotating the ellipsoid had little effect on the indicated distance (C). The settings with the light on showed the same systematic deviation from being spherical that was present with the light off, but the points were closer to the veridical value (B). Turning on the light increased the slope of indicated distance versus simulated distance (D).

the thin curve indicates the kind of errors that would arise from misjudging the distance of a 20% larger ball. The distance the subject indicated manually (D) deviated considerably from the simulated distance (see Foley & Held, 1972 for similar results), which is consistent with the notion that he misjudged the distance under these conditions, but it remains to be shown that the distance of the target was misjudged, rather than the position of the unseen hand.

Fig. 5 shows what happened to this subject's settings when either information about shape (rotating ellipsoid) or about distance (light on) was added. Adding information about shape influenced the set depth—making the shape more or less spherical—without changing the set width and height or the indicated distance. This can be seen in Fig. 5A and C: the set depths for the rotating ellipsoids (open circles) appear to be shifted vertically with respect to the set depths for static ellipsoids (filled circles) to obtain values that correspond with the set widths (diagonal line), while the range of set widths and the slope of indicated versus simulated distance did not change.

Turning on the light increased the dependency of the indicated distance on the simulated distance, albeit modestly (D; the slope increased from 0.5 to 0.66), and

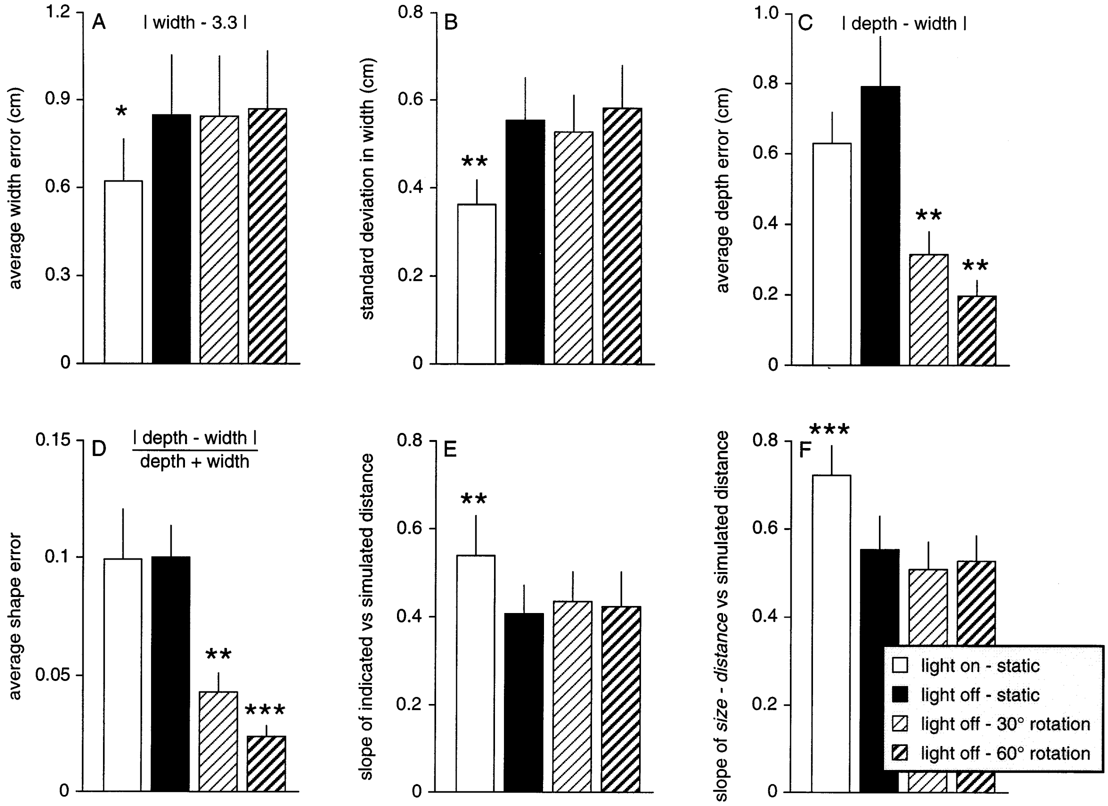

Fig. 6. Group means and between subject standard errors of various measures of the misperception of size (A, B), shape (C, D) and distance (E, F). The stars indicate that the values were significantly different from those of the light off-static condition (paired *t*-tests: * $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$).

brought both the set depth and the set width and height closer to the veridical setting for a tennis ball (the intersection of the two dotted lines; B). The set width and depth still clustered in approximately the same manner below the curve representing the errors one would expect from misjudgements of distance, suggesting that misjudgements of distance constitute a major source of errors even when the light is on.

In order to evaluate the performance of our group of subjects as a whole, we calculated several measures of errors in perceived size, shape and distance. The average value for each measure and condition is shown in Fig. 6. Error bars represent standard errors across subjects. Values that differ significantly from those of the most limited condition (light off; static) are indicated by stars.

Perceived size was expressed as the set width and height of the simulated ellipsoid (the width and height were always the same, so we will henceforth refer to this measure as the width). This choice was inspired by our wanting the measure of size to be independent of perceived depth, only requiring a combination of retinal extent with an estimate of viewing distance. Errors in set width were expressed both relative to the veridical width of a tennis ball (A), and—to allow for systematic misjudgements of the size of a tennis ball—relative to the average width that was set (B). Turning on the light decreased both the deviation from veridicality (A) and the variability between settings (B). Rotating the ellipsoid had no effect on either measure.

The most straightforward measure of the error in perceived shape is the difference between the width and the depth (C). For a sphere the two should be equal. Rotating the ellipsoid clearly reduced this 'depth error'. Turning on the light did appear—on average—to reduce the depth error, but the effect was not consistent across subjects (and was therefore not statistically significant). One drawback of using the difference between depth and width as a measure of the error in perceived shape is that it is not independent of size: a twice as large ellipsoid of the same global shape has a twice as large depth error. Dividing this depth error by the sum of depth and width provides us with a 'shape error' that is independent of size (D). Turning on the light clearly did not influence this second measure of the perceived shape. The influence of rotating the ellipsoid is still evident, as was to be expected because rotating the ellipsoid did not influence the set size.

In contrast to perceived size and shape, which the subjects were trying to keep constant across trials (to match the tennis ball), perceived distance was free to change with the simulated distance. A simple measure

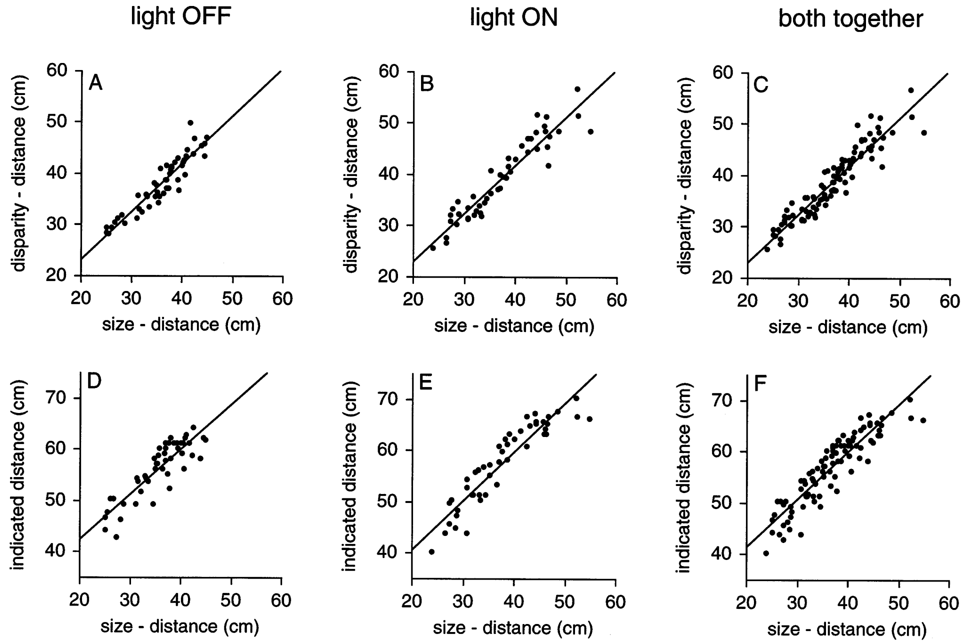

Fig. 7. Example of plots of disparity-distance versus size-distance (A–C) and of indicated distance versus size-distance (D–F) for the data from Fig. 5B and D. Note that although the range is larger with the light on (B, E) than with it off (A, D), the relationship between the measures does not change. Regression lines were fit to the data in each condition separately, as well as to both together (C, F).

of the extent to which this occurred is the slope of the linear regression of the indicated distance as a function of the simulated distance (E; for examples see Fig. 5C and D). In the most limited condition (light off; static) the average slope was about 0.4. Turning on the light increased the slope, but it remained well below 1. This may at least partly be due to the remaining conflict with accommodation and to the angular size of the triangles not decreasing with the simulated distance. Rotating the ellipsoid had no effect.

An indirect measure of perceived distance can also be derived from the set width. We defined 'size-distance' as the distance at which the set retinal size would be correct for a tennis ball (see methods). The slope of size-distance as a function of simulated distance also increased when the light was turned on (F), but did not change when the ellipsoid was rotating. The slope itself was larger in this study than in our previous study (van Damme & Brenner, 1997), possibly due to the target being nearer to the observer in the present study.

In summary, rotating the ellipsoid improved judgements of shape considerably (C, D; as in Johnston, Cumming & Landy, 1994), but did not influence the perceived size or distance. In contrast, turning on the light improved judgements of size and distance, but did not improve the perceived shape. The latter appears to be inconsistent with our earlier conclusion that the errors in the set shape are due to misjudging the distance (see Fig. 4C and van Damme & Brenner, 1997). If the errors arise from the distance being misjudged, turning on the light is expected to improve all three judgements. However, it is evident from Fig. 6 that the choice of a measure of shape could influence our conclusion on this matter.

A more direct way of determining whether errors in perceived size, shape and distance are all related to a single misjudgement of distance is by examining whether the errors in the three measures vary consistently between individual trials. In order to do so we first determined the distance that would account for each setting (see methods). We then fit lines to plots of 'disparity-distance' versus 'size-distance' and 'indicated distance' versus 'size-distance'. This was done separately for trials with the light on and off (static ellipsoid), as well as for both together. Fig. 7 shows the data for the same subject as in Figs. 4 and 5. The range of values is larger with the light on (B, E) than when they were off (A, D)—in accordance with the larger slopes in Fig. 6E and F—but the relationships themselves are very similar. The 'disparity-distance' and the 'size-distance' (A-C) are even almost, but not quite, the same. The small but systematic difference (the intercept is not quite zero, and the slope is slightly smaller than one) could be due to the subject having misjudged the size of the reference. The indicated distance is clearly different (D–F; note the different values on the axes), but is again very similar with the light on, as with it off.

In order to quantify the impression that the relationship had not changed when we turned on the light we fit separate lines to the relationships with the light on (A, D) and off (B, E), as well as a single line to both together (C, F), and compared the residual variability.

We did this for each subject. The average magnitudes of the residual variabilities are shown in Fig. 8. The strongest support for the notion that the same measure of distance was involved in all three judgements, is that the residual variability did not decrease when the lights were turned on (improving the judgement of distance); if anything, the variability increased. For the statistical analysis, we examined whether the residual variability was significantly larger when the conditions were combined, than when they were treated separately. This was done by comparing the value for both together, with the average of the values with the light on and off. Fig. 8 already suggests that this makes no difference. In fact, the largest individual F-value we found was $F_{98,96} = 1.05$ (far below the value of about 1.4 which indicates a 5% level of significance). Thus we find no evidence that the relationship between the measures has changed, which is consistent with the same distance measure being used to judge the position as for scaling retinal extent and retinal disparities.

4. Discussion

As was to be expected for single objects in the dark, subjects misjudged distances considerably. When forced to rely on retinal extent and relative disparities they also misjudged the objects' sizes and shapes. The regularities in these misjudgements suggested that they were largely due to a common incorrect estimate of distance. This was confirmed by calculating the measure of distance that would account for each misjudgement. The correlations between these three 'distances' did not change when the environment was illuminated, as predicted by the use of a single measure of distance, rather than three similar but independent measures.

Assuming that the same measure of distance did indeed contribute to all three judgements, and that errors in this measure were largely responsible for the misjudgements, we are now in a position to interpret the influence of rotating the ellipsoid. Rotating the ellipsoid and simulating a non-uniform illumination adds information about the object's shape. As recovering depth from motion parallax and shading scales differently with viewing distance than does recovering depth from binocular disparities, misjudgements of distance introduce a conflict between the depth cues. One way in which our visual system could deal with such conflicts is by some form of averaging of the resulting measures of depth (see next section). However, the conflict itself could be used as an additional source of information about distance (Johnston, Cumming & Parker, 1993; Johnston, Cumming & Landy, 1994; Frisby, Buckley, Wishart, Porrill, Gårding & Mayhew 1995), because any attempt to decrease the discrepancy between the cues will involve improving the estimate of distance. If so, then considering that the same measure of distance is used for all three judgements, improving judgements of shape should lead to better judgements of size and distance as well. Adding shading and motion parallax improved the judgements of shape considerably. However, neither the manual estimates of the objects' distances nor the set sizes were influenced. Thus our visual system does not appear to strive for consistency between percepts. This is consistent with a modular approach to vision, whereby different attributes (e.g. shape and position) are determined as independently as possible, rather than being combined in a search for a single consistent interpretation of the scene (as for instance proposed in Bülthoff, 1991).

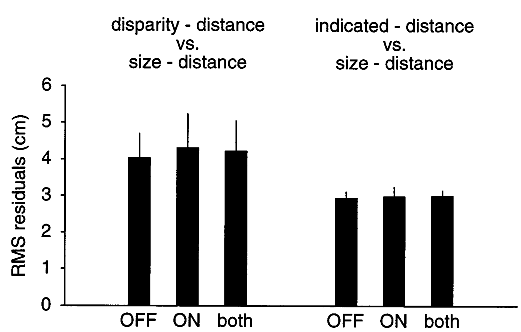

Fig. 8. Group means (and between subject standard errors) of root-mean-squares of residuals of fits as shown in Fig. 7. Turning on the light clearly did not decrease the variability. Moreover, the average of the values with the light on and off is very close to the value for both together, indicating that the fit lines are very similar.

4.1. Cue interactions

The fact that the perceived shape improved while the perceived distance did not has consequences for our understanding of how cues are combined. One advantage of simultaneously having several kinds of information (cues) about a single aspect of our visual surrounding is that this can help overcome the limitations of the individual cues. Each cue has its own limitations and requirements (e.g. motion parallax requires that either the object or the observer move), and is therefore most suitable under certain conditions. Moreover, each cue is based on certain assumptions (e.g. that the surface texture is homogeneous), and is best suited to provide a certain kind of information (e.g. positions; orientations). It makes sense, therefore, to modify the choice of cues, or the weight assigned to cues, in accordance with the conditions and the desired judgement (Gillam 1968; Young, Landy & Maloney, 1993; Landy, Maloney, Johnston & Young, 1995). Additional weight may be given to cues that provide consistent values (which may account for the reduced thresholds in Bradshaw & Rogers (1996) that were not found in Cornilleau-Pérès & Droulez (1993)), because two sources of information are unlikely to erroneously come up with exactly the same error (Todd, 1985). Moreover, conflicts can sometimes better be resolved by one cue receiving no weight at all (Bülthoff & Mallot, 1988; Cornilleau-Pérès, Marin & Droulez, 1996), than by opting for a compromise, because a compromise will often involve rejecting many assumptions, even when only one is violated. Finally, the interaction may take place before the cue is used for the judgement at hand, as when segregation by motion helps in finding corresponding points in the two eyes (Tittle & Braunstein, 1993).

In the present study, subjects had to deal with many conflicts between cues. The most obvious is that the required accommodation was fixed throughout the experiment, but there were also more subtle inconsistencies. For instance, in the most limited condition, all the simulated triangles had the same luminance. Although this may appear to abolish shape from shading, it does not necessarily do so, because the absence of changes in luminance can be considered as an indication that the surface is flat. The data in Fig. 4C show that neither this conflicting shading cue nor the consistent texture cue had much influence on the settings, because the curve that the settings appear to follow is based on the assumption that the set depth was based exclusively on relative disparities.

Rotating the ellipsoid made the conflict more evident, giving the subject the opportunity to adjust his or her measure of the ellipsoid's distance while seeking for a consistent interpretation of the shape cues. However, there was no indication that subjects changed their judgements of the ellipsoid's distance to increase the coherence between ocular convergence, shading, motion parallax, retinal extent and relative disparity. Instead, information from stereopsis appeared to have been abandoned once more reliable shape cues were added. We were surprised that the conflict was never detected; i.e. that subjects never reported perceiving a non rigid object. We do not consider this as evidence for an adjustment of the measure of viewing distance, however, because Hogervorst, Kappers & Koenderink (1997) predicted, on the basis of a study in which they explicitly examined how well subjects can distinguish between rigidly and non rigidly moving objects, that non rigidly moving objects will often erroneously be seen to move rigidly.

The smaller weight assigned to stereopsis in the present study, than in at least some conditions of Johnston, Cumming & Landy's (1994) study, is probably due to the larger amplitude of rotation in the present study, and to support for the depth specified by motion from shading, texture, and changes in contour. The relative weights assigned to motion and stereopsis are known to depend on many factors, such as the viewing distance (Johnston, Cumming & Landy, 1994), and to switch from dominance of motion to dominance of stereopsis if critical aspects of the display or of the task are changed (Norman & Todd, 1995; Turner, Braunstein & Andersen, 1997).

It is less easy to explain why we did not find that motion and stereopsis interact to improve the judgement of distance, whereas the data of Johnston, Cumming & Landy (1994) does suggest such an interaction. Apparently either the extent of this interaction also depends on the experimental conditions, or an assumption that is made in one of the studies is incorrect. For instance, we could be mistaken in our conclusion that the same measure of distance is used for all three judgements, in which case the improvement could remain restricted to shape judgements. Alternatively, their assumption that the benefit of combining motion and stereopsis involves a change in the measure of viewing distance may be incorrect; there may be some other reason for the motion cue being less effective when presented monocularly (e.g. more weight being assigned to texture when motion does not provide enough information. Such changes in weights could explain why the set shape from two frame motion depended on the viewing distance in Johnston, Cumming and Landy's study). A recent report by Econopouly & Landy (1995) that rotating an object can also improve the perceived shape of a second, static object, appears to support the first suggestion (that the improvement to the measure of distance in the shape module does not transfer to other judgements), but as the two objects were always at the same distance, subjects may have had a tendency to match the disparity of the static object to that of the moving one, because they should be the same for objects of the same shape and size.

The separation of vision into independent modules (Marr, 1982; Zeki & Shipp, 1988; Livingstone & Hubel, 1988; Landy, Maloney, Johnston & Young, 1995), and the ability to ignore information within each module, has the advantage that it allows one to use information that is specifically suited for extracting the attribute of interest, without having to consider the implications for all other attributes, and without even having to transpose all the information to one or several common representations (Landy, Maloney, Johnston & Young, 1995). As a consequence, however, the outcomes will not always be mutually consistent.

4.2. Size distance invariance

Although the manually indicated distance in Fig. 5D–F is evidently quite different from that used to judge the object's size, the fact that the relationship did not change when more information about distance was added suggests that a single measure of distance was employed in both judgements, so that the difference must be due to systematic errors in some other information that is involved. For the size-distance this could be the measure of retinal extent, or the subject's notion of the size of a tennis ball. For the manually indicated distance this could be the kinesthetic information about the position of the hand. Systematic errors in any or several of these measures would result in systematic differences between the manually indicated distance and the size-distance. Our data therefore appear to support the size distance invariance hypothesis by showing that the measure of distance that determines the perceived distance also determines the perceived size.

However, we found that adding information about shape that does not rely on a measure of distance can influence the perceived shape without affecting the perceived size or distance. Similarly, one can conceive of cues that provide information about perceived size that are independent of distance, such as familiarity with an object. When such cues are used, the perceived size is known to change independently of the perceived distance (Kilpatrick & Ittelson, 1953; Gogel, 1990). Thus, although there is a single measure of distance, the extent to which the perceived size will be consistent with this measure depends on the extent to which the perceived size is based on information that requires a measure of distance (i.e. on retinal extent). In conclusion, the invariance that is observed under some conditions, such as those of the present study, is probably the result of the same measure of distance being used whenever such a measure is required, rather than from consistency between the measures being explicitly imposed.

4.3. Separate pathways for perception and action

The measure of distance that was used to interpret the retinal image size and to interpret relative disparities also determined the manually indicated distance. This not only suggests that a single measure of distance is used throughout perception, but also that the same measure is used when making visually guided movements. It has repeatedly been proposed that there are separate pathways for perception and action (Bridgeman, Kirch & Spirling, 1981; Goodale & Milner, 1992). We have previously questioned this suggestion by showing that the difference lay in the visual judgement that was involved, rather than in the kind of response that was required (i.e. by reporting on a percept or by direct manual action; Smeets & Brenner, 1995; Brenner & Smeets, 1996b). The present results are consistent with our view of independent modules for different aspects of vision (e.g. size, shape, position, motion, colour), each contributing to both perception and action.

5. Conclusion

Judgements of size, shape and distance take place in separate modules. A considerable amount of consistency between the judgements is normally maintained by the consistency in the input from the environment, and by the common use of certain measures (in our study an estimate of viewing distance) when interpreting different kinds of information (e.g. retinal extent; retinal disparities). Inconsistencies between percepts could indicate that one or more judgements is incorrect, but this information does not appear to be used to improve the judgements.

References

Arend, L., & Reeves, A. (1986). Simultaneous color constancy. *Journal of the Optical Society of America A*, *3*, 1743–1751.

van den Berg, A. V. (1996). Judgements of heading. *Vision Research*, *36*, 2337–2350.

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

Biguer, B., Donaldson, I. M. L., Hein, A., & Jeannerod, M. (1988). Neck muscle vibration modifies the representation of visual motion and direction in man. *Brain*, *111*, 1405–1424.

Bradshaw, M. F., & Rogers, B. J. (1996). The interaction of binocular disparity and motion parallax in the computation of depth. *Vision Research*, *36*, 3457–3468.

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

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

Brenner, E., & Smeets, J. B. J. (1996a). Hitting moving targets: Co-operative control of 'when' and 'where'. *Human movement science*, *15*, 39–53.

Brenner, E., & Smeets, J. B. J. (1996b). Size illusion influences how we lift but not how we grasp an object. *Experimental Brain Research*, *111*, 473–476.

Bridgeman, B., Kirch, M., & Spirling, A. (1981). Segregation of cognitive and motor aspects of visual function using induced motion. *Perception and Psychophysics*, *29*, 336–342.

Bruno, N., & Cutting, J. E. (1988). Minimodularity and the perception of layout. *Journal of Experimental Psychology*: *General*, *117*, 161–170.

Bülthoff, H. H. (1991). Shape from X: psychophysics and computation. In M. S. Landy, & J. A. Movshon, *Computational models of visual processing*. Boston, MA: MIT press, 305–330.

Bülthoff, H. H., & Mallot, H. A. (1988). Integration of depth modules: stereo and shading. *Journal of the Optical Society of America A*, *5*, 1749–1758.

Collett, T. S., Schwarz, U., & Sobel, E. C. (1991). The interaction of oculomotor cues and stimulus size in stereoscopic depth constancy. *Perception*, *20*, 733–754.

Collewijn, H., & Erkelens, C. J. (1990). Binocular eye movements and the perception of depth. In E. Kowler, *Eye movements and their role in visual and cognitive processes*. New York, NY: Elsevier.

Cornelissen, F. W., & Brenner, E. (1995). Simultaneous colour constancy revisited: an analysis of viewing strategies. *Vision Research*, *35*, 2431–2448.

Cornilleau-Pérès, V., & Droulez, J. (1993). Stereo-motion cooperation and the use of motion disparity in the visual perception of 3-D structure. *Perception and Psychophysics*, *54*, 223–239.

Cornilleau-Pérès, V., Marin, E., & Droulez, J. (1996). The dominance of static depth cues over motion parallax in the perception of surface orientation. *Perception*, *25*(Suppl.), 40.

Cumming, B. G., Johnston, E. B., & Parker, A. J. (1991). Vertical disparities and perception of three-dimensional shape. *Nature*, *349*, 411–413.

Cutting, J. E., & Vishton, P. M. (1995). Perceiving layout and knowing distances: the integration, relative potency, and contextual use of different information about depth. In W. Epstein, & S. Rogers, *Perception of space and motion*. Orlando, FL: Academic press, 69–117.

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.

Foley, J. M., & Held, R. (1972). Visually directed pointing as a function of target distance, direction, and available cues. *Perception and Psychophysics*, *12*, 263–268.

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.

Gogel, W. C. (1961). Convergence as a cue to the perceived distance of objects in a binocular configuration. *Journal of Psychology*, *52*, 303–315.

Gogel, W. C. (1990). A theory of phenomenal geometry and its applications. *Perception and Psychophysics*, *48*, 105–123.

Gogel, W. C., & Tietz, J. D. (1973). Absolute motion parallax and the specific distance tendency. *Perception and Psychophysics*, *13*, 284–292.

Goodale, M. A., & Milner, A. D. (1992). Separate visual pathways for perception and action. *Trends in Neuroscience*, *15*, 20–25.

Hogervorst, M. A., Kappers, A. M., & Koenderink, J. J. (1997). Monocular discrimination of rigidly and nonrigidly moving objects. *Perception and Psychophysics*, *57*, 1266–1279.

Jeannerod, M. (1981). Intersegmental coordination during reaching at natural visual objects. In J. Long, & A. Baddeley, *Attention and performance*, Vol. IX. Hillsdale: Erlbaum, 153–168.

Johnston, E. B. (1991). Systematic distortions of shape from stereopsis. *Vision Research*, *31*, 1351–1360.

Johnston, E. B., Cumming, B. G., & Parker, A. J. (1993). Integration of depth modules: stereopsis and texture. *Vision Research*, *33*, 813–826.

Johnston, E. B., Cumming, B. G., & Landy, M. S. (1994). Integration of stereopsis and motion shape cues. *Vision Research*, *34*, 2259–2275.

Kilpatrick, F. P., & Ittelson, W. H. (1953). The size-distance invariance hypothesis. *The Psychological Review*, *60*, 223–231.

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 Fransisco, CA: W.H. Freeman.

McCready, D. (1985). On size, distance, and visual angle perception. *Perception and Psychophysics*, *37*, 323–334.

Norman, J. F., & Todd, J. T. (1995). The perception of 3-D structure from contradictory optical patterns. *Perception and Psychophysics*, *57*, 826–834.

Paulignan, Y., & Jeannerod, M. (1996). Prehension Movements: the visuomotor channels hypothesis revisited. In A. M. Wing, P. Haggard, & R. Flanagan, *Hand and brain*. Orlando, FL: Academic Press, 265–282.

Rogers, B. J., & Bradshaw, M. F. (1995). Binocular judgements of depth, size, shape and absolute distance: is the same 'd' used for all judgements? *Investigative Ophthalmology and Visual Science*, *36/4*, 230.

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.

Sedgwick, H. A. (1986). Space perception. In K. R. Boff, L. Kaufman, & J. P. Thomas, *Handbook of Perception and Human Performance*; *Sensory Processes and Perception*. New York, NY: Wiley-Interscience, 1–57.

Smeets, J. B. J., & Brenner, E. (1995). Perception and action based on the same visual information: distinction between position and velocity. *Journal of Experimental Psychology*: *Human Perception and Performance*, *21*, 19–31.

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.

Todd, J. T. (1985). Perception of structure from motion: is projective correspondence of moving elements a necessary condition? *Journal of Experimental Psychology*: *Human Perception and Performance*, *11*, 689–710.

Turner, J., Braunstein, M. L., & Andersen, G. J. (1997). Relationship between binocular disparity and motion parallax in surface detection. *Perception and Psychophysics*, *59*, 370–380.

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.