Vision Res. Vol. 28, No. 3, pp. 371–386, 1988
Printed in Great Britain

0042-6989/88 $3.00 + 0.00
Pergamon Press plc

INTEGRATING STEREOPSIS WITH MONOCULAR INTERPRETATIONS OF PLANAR SURFACES*

KENT A. STEVENS and ALLEN BROOKES
Department of Computer Science, University of Oregon, Eugene, OR 97403, U.S.A.

(*Received* 10 *November* 1986; *in final revised form* 11 *September* 1987)

Abstract—Experiments are reported that involved spatial judgments of planar surfaces that had contradictory stereo and monocular information. Tasks included comparing the relative depths of two points on the depicted surface and judging the surface's apparent spatial orientation. It was found that for planar surfaces the 3D perception was dominated by the monocular interpretation, despite the strongly contradictory stereo information. We propose that stereo information is effectively integrated only where the surface exhibits curvature features or edge discontinuities, i.e. where the second spatial derivatives of disparity are nonzero. Planar surfaces induce constant gradients of disparity and are thus effectively featureless to stereopsis. Further observations are reported regarding nonplanar surfaces, where contradictory monocular information can still be effectively rivalrous with that suggested stereoscopically.

Stereopsis Binocular vision Depth perception

INTRODUCTION

How does stereopsis constrain the perceived 3D shape and spatial orientation of static surfaces? The most plausible answer, seemingly, would be in terms of distance information determined from disparity at points across the surface. Stereopsis is generally expected to provide 3D distance information, specifically range and relative depth across visible surfaces, as derived from horizontal (and possibly vertical) retinal disparities given geometric parameters such as the angles of gaze and convergence (Mayhew, 1982; Longuet-Higgins, 1982a, b; Prazdny, 1983). There is much psychophysical evidence to support the view that stereopsis provides distance information. Stereopsis allows accurate judgments of absolute distance out to at least 2 m (e.g. Wallach and Zuckerman, 1963; Ritter, 1977, 1979; Morrison and Whiteside, 1984), and, within that range, distance intervals are accurately perceived from disparity intervals (so-called "stereo depth constancy", see Ono and Comerford, 1977; Wallach *et al.*, 1979). It therefore seems reasonable to conclude that binocular vision in natural circumstances results in more-or-less complete and accurate 3D mapping of the surfaces in the immediate surrounds. But it is not clear how that 3D information might be combined with that derived monocularly.

Compared to stereopsis, the monocular "depth cues" in a static image provide much weaker and less precise 3D information†. Strongly restrictive assumptions are required to interpret cues such as shading, texture gradients, and monocular configurations such as in Fig. 1 (Stevens, 1981a, b, 1984). In comparison to the sound geometrical basis for determining absolute and relative distances from stereo disparity, one would expect stereopsis to dominate over the less reliable monocular information. This study and others, however, suggest the contrary: monocular configurations often dominate the resulting 3D interpretation over stereopsis, even in the near range where stereopsis is most accurate.

To be sure, binocular vision generally yields more accurate 3D judgments than monocular vision based on linear perspective, texture, shading, and so forth (e.g. Smith and Smith, 1957, 1961; Smith, 1965). Contradictory results were reported by Youngs (1976), however, where

*Supported by Office of Naval Research Contract N00014-K-84-0533.

†Monocular depth cues, despite their name, are primarily sources of information about local surface orientation (the orientation of surface patches relative to the line of sight) and of shape (surface curvature as well as the intrinsic geometry of the surface) and only in a weaker sense able to deliver distance information, either relative or absolute (Marr, 1982; Stevens, 1983b). That is, monocularly there is more reliable information about surface shape features and orientation than of distance *per se.*

Fig. 1. Monocular configurations that evoke definite 3D interpretations.

stereo disparity had no significant effect on apparent slant (of planar stimuli). Youngs (1976) questioned "why the disparity coding fails so miserably" in those experiments. Stereopsis is particularly weak in the presence of a strong contradictory monocular interpretation, such as presented in reversed-disparity stereograms of a face or a street scene (Wheatstone, 1852; Schriever, 1925; Gregory, 1970; Yellott and Kaiwi, 1979), or by Hochberg's striking Necker cube stereogram (see Julesz, 1971, p. 163), wherein a cube at constant retinal disparity readily reverses in depth.

We performed a series of experiments to attempt to determine what role stereopsis plays in the presence of contradictory monocular information. Experiment 1 concerned whether stereopsis could be used to effectively contradict the monocular interpretation of oblique intersections as foreshortened right angles, when the intersections were actually not perpendicular in 3D. We used stimuli similar to the planar grid in Fig. 1, and found stereopsis remarkably impotent in influencing the perceived orientation and 3D configuration. Experiment 2 similarly examined relative depth judgements in displays with conflicting stereo and monocular information. Given a simple pair of stereo points, that with the greater (more positive) disparity is seen as relatively farther. But if these points are embedded in a continuous 3D surface, and if the monocular interpretation suggests an alternative relative depth between the two points, that monocular interpretation governed the judgement in our experiment. Experiment 3 similarly examined whether a conflicting disparity gradient influenced the monocularly interpreted surface orientation.

We recognized a common theme: our stimuli, although rich in terms of stereo information, consisted of planar surfaces in 3D. Examination of control stimuli convinced us that sufficient stereo information was available, rather it appeared that stereo disparities across a planar surface were simply not effectively analyzed in 3D. More formally, we hypothesized that stereopsis extracts 3D surface information only where the second spatial derivatives of disparity are nonzero, corresponding to loci where the surface is curved, creased, or discontinuous. Experiment 4 directly examined planar versus nonplanar stereo stimuli, with and without competing monocular interpretations. The results further support this hypothesis. (And reviewing earlier studies, we observed that where stereopsis was particularly ineffective against conflicting monocular information, those studies involved planar surfaces.)

An adequate explanation must address two issues: the computation of depth from disparity and the integration of stereo and monocular 3D information. We will argue that depth is derived from disparity only where the surface exhibits continuous curvature or sharp discontinuities. But we suggest that depth, the apparent variation in surface relief, is reconstructed from multiple sources of evidence about surface topography. That is, surface shape is first analyzed in terms of sharp edges and creases, smooth folds, indentations, and so forth, from both binocular and monocular sources. The depth one experiences is a consequence of how this information is interpreted and reconciled. Depending on how the monocular information is interpreted, radically different depth distributions might be experienced. This is quite distinct from the notion that depth (and slant) is derived directly from stereo disparity (and its gradient).

EXPERIMENTS

Experiment 1: Interpretation of Perpendicular Intersections

Observers tend to interpret monocular images of oblique intersections as right-angle intersections in 3D (Attneave and Frost, 1969; Perkins, 1972; Shepard, 1981; Stevens, 1983a). In an earlier experiment, Stevens (1983a) found that subjects perceive such stimuli (e.g. a cross or a parallelogram) as lying on a plane oriented in 3D. Subjects could reliably visualize the orientation of that plane, and judge whether a line segment, superimposed on the monocular stimulus at a given image orientation, corresponded to the visualized normal to the plane. Moreover, apparent tilt (direction of slant) agreed closely with that predicted by assuming that the stimulus image corresponded to a right angle in 3D. In the present experiment we used similar cross and grid stimuli, but now projected stereoscopically, in order to examine whether the available stereo information would permit observers to distinguish the true 3D configuration.

Method

Apparatus. Stereo pairs were presented by a Wheatstone-style stereoscope using a pair of optically flat front-surfaced mirrors and two Tektronix 634 monochrome displays (flat 9 × 12 cm screens, 1100 line resolution, and less than 0.5% geometric distortion). The optic path from monitor screen to observer was 38 cm, and the two paths converged at total angle of 9.8° (providing consistent accommodation and vergence for a 65 mm interpupillary separation). Circular apertures allowed a 6.4° radius field of view. The stimuli consisted of luminous lines against a dark background. The stereograms were generated dynamically by a Symbolics 3670 Lisp Machine; the monochrome monitors projecting the left and right images were driven independently by separate channels of a color frame buffer.

To generate a stereo pair, 2D projections were computed from left and right vantage points that differed by the 9.8° convergence angle. The images could be generated in either perspective or orthographic projection. In the perspective case (used in Experiments 2 and 3) the projection was computed as if the surface were physically situated 38 cm from the viewer; for the orthographic case (Experiments 1 and 4) the viewing distance was 100-fold further with the image scaled accordingly so as to subtend the same visual angle as in the perspective case. All computed stereo disparities were distributed equally to the two half-images, corresponding with a frontal, foveal viewpoint with symmetrical convergence of the two eyes.

Stimuli. Two types of orthographic stimuli were presented stereoscopically: a pair of crossing lines and a 5 × 5 grid of lines. The angle of intersection was either 90° (Fig. 2) or skewed 15, 30 or 45° from the perpendicular (Fig. 3). The grid became an increasingly racked parallelogram with increasing skew angle. Monocularly, varying skew angle would imply different spatial orientations; stereoscopically the spatial orientation should remain constant and only the intersection angle should appear to vary. The intention was to place a compelling monocular* impression of perpendicularity in opposition to contradictory stereo information. Note that orthographic projection was used to avoid a monocular cue to skew angle provided by perspective distortion to the skewed grid.

The stimuli were specified by three spatial parameters relative to the plane containing the grid or cross. The orientation of the plane in stereo was defined by its slant (the angle between the normal to the plane and the line of sight) and tilt (the direction to which the normal would project, i.e. the direction of slant). The third parameter specified the angular orientation of the grid or cross on the slanted plane (a rotation about the normal to the plane). The slant was held constant at 65°. Three angles of tilt and two angular orientations were used to provide six visually distinct perspectives of the grid and cross stimuli for each of the four skew angles—see (Stevens, 1983a) for similar cross and grid experiments in which the accuracy of apparent tilt judgments was found to be substantially independent of the choice of tilt angle.

Procedure

Ten graduate students participated as paid subjects; all had good stereo vision and were naive to the purposes of the experiment. The subjects were shown example stimuli and explained that they would see crosses and grids oriented at a slant relative to the observer and that the 3D intersections would sometimes be right angles and at other times skewed (the notion of a skewed intersection was reinforced with a physical demonstration). They were to make force-choice judgments of whether the intersection was perpendicular in 3D or not (referred to as the P judgment, made by depress-

*Here we refer to the fused binocular image as a 2D projection, in Julesz's (1971) sense of a "cyclopean" retina. The projection might be described geometrically as the average of the left and right half images, or the equivalent projection that would arise with a zero interpupillary separation. We will refer to the "monocular" information present in that projection, disregarding the disparity information that is present as well.

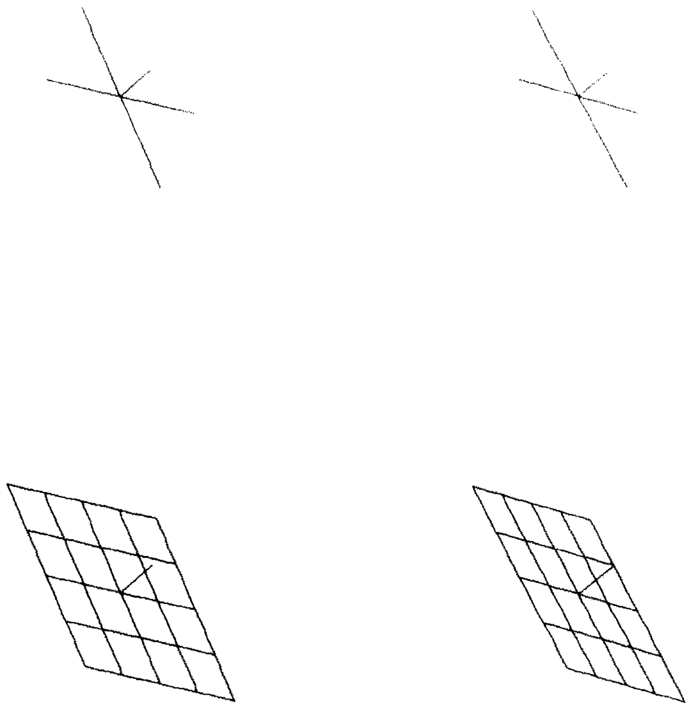


Fig. 2. Examples of cross and grid stereograms, each with 0° skew angles. Note that the normal appears to project perpendicularly to the plane defined by the cross or grid.

ing a mouse button). A positive response corresponded to lines that appeared within approximately 5° of perpendicular. Unlimited presentation time was allowed. The P judgment response initiated the addition of a stereo line segment to the stimulus that was a geometrically accurate rendition of the normal to the plane of the cross or grid. The subject made a second

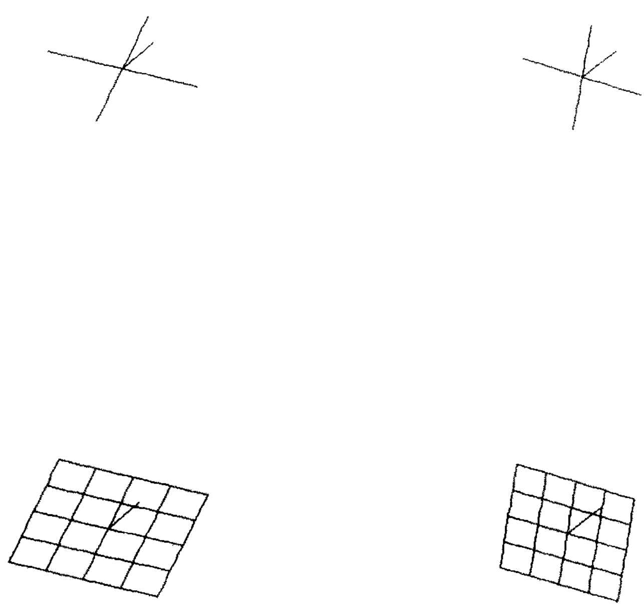
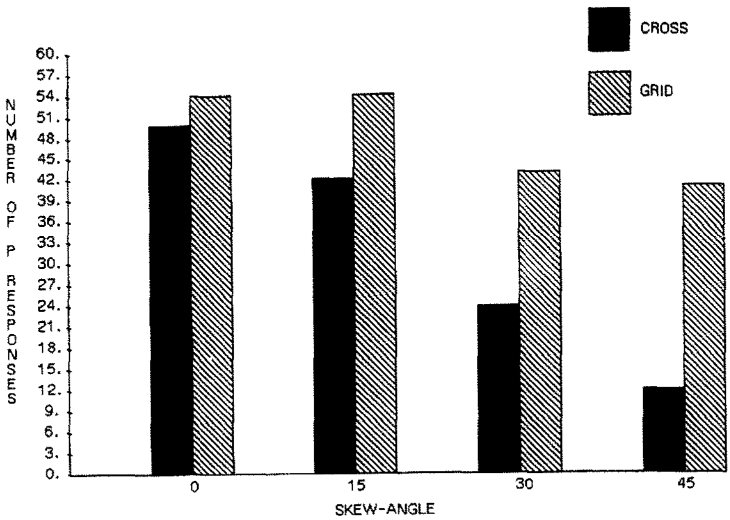
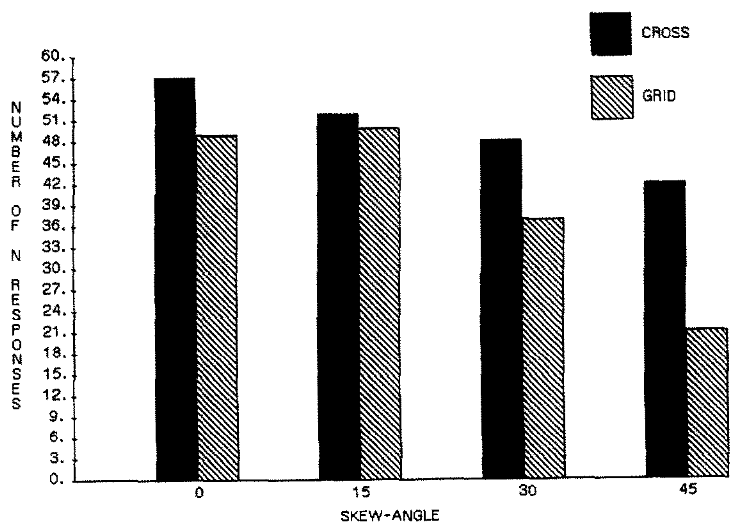


Fig. 3. Cross and grid stereograms, with identical spatial orientation as in Fig. 2, but with intersections skewed 45° from perpendicular. Note that the "normals" do not appear perpendicular to the plane of the grid or cross.



A



B

Fig. 4. Judgments of perpendicularity as a function of skew angle for cross and grid stimuli in (a); corresponding judgments of the surface normal in (b).

forced-choice response whether the line appeared to be normal (the N judgment, with the same criterion of roughly 5°).

Results and discussion

Figures 4(a) and (b) graph the number of P and N judgments as a function of skew angle for the cross and grid stimuli. For 0° skew the monocular and stereo information are both consistent with right angle intersections on a plane slanted 65°. Hence the 0° skew condition provides a baseline for the P and N judgments at greater skew. As skew angle increased, the N and P judgments for crosses and grids showed

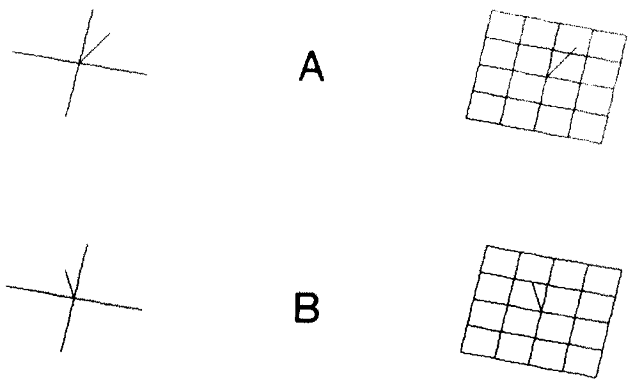


Fig. 5. In (a) the normal is correct for the monocular projection of a cross skewed 45°. In (b) the normal is correct of the monocular projection of a right angle intersection.

different, and complementary, trends. Concerning the P judgments, the grids had a greater tendency to be seen as perpendicular, and correspondingly, the displayed normals appeared increasingly incorrect as skew angle increased. The crosses were seen more vertically (i.e. according to the stereo information) although both P and N decreased with increasing skew for the crosses as well. Overall the grids were much more persistently judged on the basis of the monocular information. These trends all showed significance at $P < 0.05$ using sign tests comparing the N and P judgments for 0° and 45° skew angles.

Since the stereo projection of the normal was geometrically correct with regard to the plane containing the intersecting lines, regardless of their angle of intersection in 3D, if stereopsis had dominated the P and N judgments, the intersections would have appeared skewed for all but the 90° case and the normals would have always appeared correct. Conversely, if the judgments were based on the monocular information, the intersections would have always appeared perpendicular and the normal would have appeared incorrect except for the 90° case.

The data fell between these two alternatives: the monocular interpretation was markedly influential despite the geometrically-correct stereo information, and significantly more so for the grid than the cross. We also note that the subjects' overall ability to judge the intersection angle was not particularly sensitive (e.g. skew angles differing by 15° were barely distinguishable).* Thus the lack of precise correspondence between the N and P judgments as a function of skew angle may reflect the differences in difficulty of the two tasks.

Figure 5(a) depicts the tilt of the surface normal for a cross and grid that is skewed 45°. This figure was rendered by projecting an experimental stimulus, with the geometrically-correct surface normal, at 0° rather than 9° convergence angle. Note that the normal in Fig. 5(a) seems incorrect. Figure 5(b), which appears more appropriate, was computed by assuming the projection corresponds to a square cross or grid (see Stevens, 1983a, appendix, for formula). Figure 5(b) thus illustrates the difference between the geometrically-correct stereo interpretation of a 45° intersection, and what one would perceive if that intersection were assumed perpendicular.†

Given the richer stereo information in the grid stimulus (10 lines and 25 intersection points, compared to 2 lines and one intersection point) one might expect more accurate spatial localization of the grid than the cross. But stereopsis had a weaker role in determining both the perceived 3D orientation of the grid and the angle of intersection of the grid lines, compared to the simpler cross stimulus. There was seemingly a greater tendency to "ignore" the stereo information in the grid compared to the cross stimuli.

*We later asked two experienced observers to judge the angle of intersection for various cross stimuli and found that they could accurately estimate the true intersection angle to within 5° or so, and yet, for the correspondence grid stimuli, they repeatedly judged a 45° intersection to be skewed only 15° or so from perpendicular.

†Quantitatively, the difference in tilt amounts to 64°. The slant is also influenced by assuming the intersection is 90°. For example, the grid stereogram in Fig. 3 appears slanted much less than 65°). The computed monocular slant for Fig. 3, assuming it corresponds to a square grid, is only 38.5°.

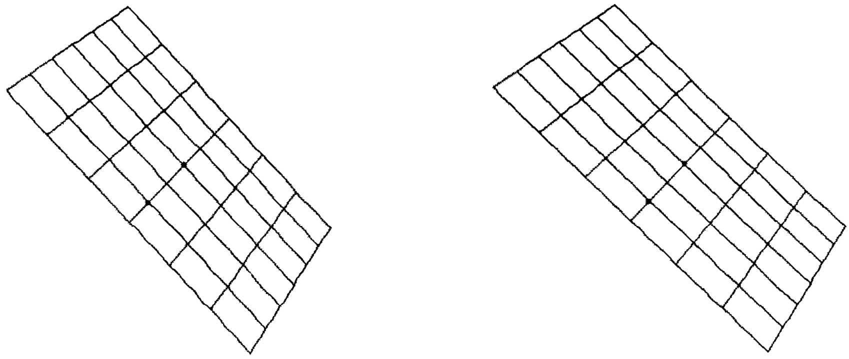


Fig. 6. Example stimulus in which subjects judged whether the given probe point was nearer than, equidistant, or further than the central reference point. The stereo disparity gradient was either consistent with, orthogonal to, or opposite from the monocularly implied distance gradient.

Experiment 2: Two-Point Relative Depth Judgments

Method

Stimuli. The optical arrangement was unchanged from Experiment 1, but we now decoupled the computation of stereo disparities from the monocular projection of the individual half-images. The aim was to examine the influence of conflicting stereo and monocular information on the judgement of the relative depth of two points on the depicted surface. The stimulus surface was a 7 × 7 square grid of lines projected in perspective, slanted 65° as in Experiment 1, and tilted either 45 or 135°.

To compute the stereogram, the screen coordinates of the two half-images were first projected according to a 0° vergence angle, which would have resulted in identical half-images, except for the introduction of horizontal disparities that were either consistent or inconsistent with the monocular projections. Four cardinal directions were defined on the stimulus surface, with north corresponding to the monocular direction of tilt (i.e. distance increased to the north on the basis of perspective). The stereo and monocular information were consistent when the stereo disparity gradient was northward. When the gradient increased to either the east or west it was orthogonal to the monocular perspective, and when to the south the stereogram had effectively reversed disparities. The surface at the central reference point always had zero disparity.

Procedure. The four subjects had participated earlier in the first experiment. The task was to judge whether a given probe point was nearer or further than, or at the same depth as a reference point located at the center of the surface. The probe point was 6° away from the reference point in one of the four cardinal directions (Fig. 6). Both probe and reference points subtended 10′ and were projected stereoscopically with disparities corresponding to points embedded in the stereo surface of the grid. There were 5 repetitions of the 32 stimuli: 2 tilts (45° and 135°), 4 probe locations (N, S, E, W), and 4 directions for the disparity gradient, in random order.

Results and discussion

Table 1 shows the sets of relative depth responses for each combination of probe

Table 1. Percentage of judgments that the probe point appeared nearer than (<), equidistant (=), or farther than (>) the central reference point. The relative depth predicted on basis of stereo disparities is in bold

	Probe location											
Direction of	N			S			E			W		
disparity gradient	<	=	>	<	=	>	<	=	>	<	=	>
North	0	0	**100**	**100**	0	0	25	**53**	22	18	**55**	27
South	**3**	12	85	92	8	**0**	8	**67**	25	22	**70**	8
East	0	**0**	100	100	**0**	0	33	42	**25**	**18**	60	22
West	0	**13**	87	87	**13**	0	**18**	60	22	22	63	**15**

Fig. 7. The disparity gradient is perpendicular to the apparent monocular gradient of distance. Subjects adjusted the monocular "normal" by rotating it in the image plane until it appeared perpendicular to the grid in 3D, i.e. to align with the surface normal.

location and disparity gradient direction. The values in boldface indicate the responses consistent with the stereo disparities. The first row serves as a control, since the direction of the stereo and monocular gradients coincided. For this case the N and S probe locations show the expected depth judgments. The E and W probe locations were generally judged equidistant, but there were also several "farther than" and "nearer than" judgments. The "equidistant" judgment turned out to be problematic. Since the two half-images were projected in perspective, points due east and west of the central reference point would have been necessarily farther than the reference point simply by the perspective projection. We thus carefully computed the E and W probe locations to be slightly south of due east and west so that, monocularly, they and the reference point were equidistant from the observer. Nonetheless it turned out rather difficult to decide whether the E, W, and reference points appeared equidistant, even with consistent stereo information, and even for highly experienced observers.

When disparity was reversed (Table 1, second row) there was an overwhelming tendency to continue to see the N point as farther, and the S point as nearer, that is, according to the monocular perspective and contrary to the stereo disparities. Some "regression to the frontal plane", is apparent, suggesting that subjects experienced a reduced impression of depth or slant in this case, as Gillam (1968) also found in reversed-disparity stereograms.

The important cases, we believe, concern disparity gradients *orthogonal* to the monocular distance gradient. Consider, for example, the case of the disparity gradient to the west and the probe point west of the reference point. The probe had positive disparity, and on that basis should have been seen as farther, but was not. The direction of the disparity gradient had no systematic effect on the depth judgments for the east and west probe locations. Overall, the apparent depth corresponded very closely with the monocular perspective, despite the contradictory stereo information.

Experiment 3: Surface Orientation Judgments

The results of Experiment 2 suggested that a disparity gradient orthogonal to the perspective distance gradient had negligible influence on the relative depths of two points on the surface. Experiment 3 pursued this result in terms of the effect of a competing disparity gradient on apparent tilt—see method in (Stevens, 1983a). Subjects adjusted a needle to appear perpendicular to the apparent plane of the grid. If the orthogonal disparity gradient had an effect, we would expect the needle to lean in the direction of the stereo gradient, an effect analogous to the vector sum of the monocular and stereo interpretations.

Method

Stimuli. Stereograms were constructed for which the stereo information corresponded to a surface whose 3D orientation was precisely orthogonal to that depicted monocularly. The stimulus surface was a 5×5 square grid of lines projected in perspective, slanted either 35 or 70° and tilted either 40 or 140°. The disparities corresponded to a slanted plane whose tilt was $\pm 90°$ away from the monocular tilt. The monocular cue implied depth increasing to the north while the stereo information implied depth increasing to either the east or west, depending on the polarity of the disparity gradient.

Procedure. Three subjects were used; all had previous experience in the experimental series. A grid surface was presented for one second prior to superimposing a rotatable line segment that had one endpoint fixed at the center of the grid. The "needle" was presented monocularly, to the

Table 2. Mean surface tilt judgments (and standard deviations) with monocular normal

Slant	Tilt	Disparity gradient to west	Disparity gradient to east
35.0	40.0	50.5 (5.3)	49.2 (7.5)
35.0	140.0	142.7 (4.5)	141.2 (4.5)
70.0	40.0	46.8 (2.2)	44.0 (4.0)
70.0	140.0	139.8 (3.5)	138.7 (3.6)

dominant eye only (see Fig. 7). Subjects stepped the needle in tilt by $\pm 2.5°$ increments until it pointed in the direction of the surface normal. The needle appeared to emerge from the surface and to pivot in 3D about the fixed end, despite only rotating in the image plane. Unlimited time was permitted per trial. Tilt data was recorded for 5 trials of each of 8 conditions (four monocular surface orientations times two directions for the stereo disparity gradients).

Apparent slant was also recorded using a stereoscopic needle that could be stepped in both slant and tilt. The three subjects were presented 5 trials per each of the eight conditions, as above.

Results and discussion

Since the disparity gradient was orthogonal to the monocular depth gradient, the apparent normal might be expected to lean in the direction of the disparity gradient, e.g. to rotate counterclockwise (increase numerically) when the disparity gradient was to the west, and clockwise when the gradient was reversed to the east. However, the data exhibited no systematic leaning in the direction of the stereo disparity gradient (see Table 2). Moreover, the apparent tilt was in reasonably close agreement with the monocularly predicted tilt. Overall, the apparent tilt seemed determined only monocularly.

Similarly, apparent slant was in close accordance with that predicted by the monocular perspective (see Table 3). This is remarkable given that the stereo disparity was constant (and zero) in that direction. The slant probe was adjusted to within one standard deviation of the slant suggested by the monocular perspective for all conditions.

Stereo disparity was constant in the direction that the monocular cues indicated increasing depth, and vice versa. With the two cues orthogonal, if they were somehow summated, one would expect the resulting apparent tilt to be influenced by the direction of the disparity gradient, but no such effect was observed. Moreover, apparent slant was in good correspondence with that predicted by the monocular perspective, despite the fact that stereo disparities were constant in that direction. This experiment thus extends the more qualitative findings of Experiments 1 and 2.

Experiment 4: Planar vs Nonplanar Stereo Disparity Distributions

In this final experiment we used line grid and random dot stereograms of planar and nonplanar surfaces to explore the importance of surface geometry on the simple two-point relative depth judgment (as in Experiment 2) in the presence and absence of competing monocular information. Our strategy was to embed a pair of stereo points in various surfaces to see to what extent the "context" influenced the apparent relative depths of these two points.

Method

Stimuli. The stimuli were grid stereograms (with lines separated by 1.9°) and random dot stereograms (Fig. 8). The horizontal disparity across the stereogram was a continuous one-dimensional function of screen position, corresponding to either a slanted plane, a Gaussian ridge, or a Gaussian-smoothed edge. These "stereo surfaces" were oriented either horizontally (*h*) or vertically (*v*). The slanted plane *v*, for example, corresponded to a plane pivoted about the vertical meridian, with disparities that varied from 0′ at the center to $\pm 51.2'$ at left and right extremes of the field of view (occluded by the optical apparatus at 6.4° eccentricity). Similarly, the Gaussian ridge function induced stereo disparities from $-37.8'$ along the ridge to 0′ in the periphery [see the horizontally oriented ridges in Fig. 8(a) and (b)]. The ridge protruded towards the viewer with half-amplitude at $\pm 1.6°$ eccentricity. The Gaussian-smoothed edge had the same space constant as the ridge. It presented a smoothed step transition from $\pm 18.9'$ at opposite edges of the field that passed through zero along the vertical or horizontal meridian [see vertical case in Fig. 8(c) and (d)].

Table 3. Mean surface slant judgments (and standard deviations) with stereoscopic normal

Slant	Tilt	Disparity gradient to west	Disparity gradient to east
35.0	40.0	36.5 (2.8)	37.5 (4.2)
35.0	140.0	33.0 (3.6)	38.8 (6.1)
70.0	40.0	68.5 (4.6)	64.7 (8.7)
70.0	140.0	65.0 (7.7)	66.5 (6.6)

Procedure. Three subjects from earlier experiments were used; all had excellent stereo vision. The task, as in Experiment 2, was to judge the depth of a probe point relative to a central reference point. The probe and reference points both subtended 10′. The probe was placed at 2.9° eccentricity either north (above), south, east (right of), or west of the reference point. The probe and reference points were both on the given stereo surface. (For the Gaussian ridge *h*, for example, the reference point had −37.8′ disparity. The probe point had 0′ disparity when north or south and −37.9′ when east or west of the reference point.) The subject indicated by mouse button whether the probe point appeared nearer, at the same depth as, or farther than the reference point. Free eye movements and unlimited observation time were allowed. The grid and dot versions of the experiment were run separately, each with 5 repetitions of the 24 conditions (six oriented disparity surfaces times four probe locations) in random order.

Results and Discussion

The relative depths of two stereo points could be determined, in principle, directly from their corresponding disparities. In a pilot experiment, where only the probe and reference points were displayed against a black background, their relative depth could be judged immediately and accurately, in accordance with their relative disparities. But when the two stereo points were embedded in a stereo surface, we found that the depth judgment depended on that surface. We conjecture that the depth judgment was mediated not directly by the relative disparities but by the perceived depth of the underlying surface. And, the perceived depth of the surface is

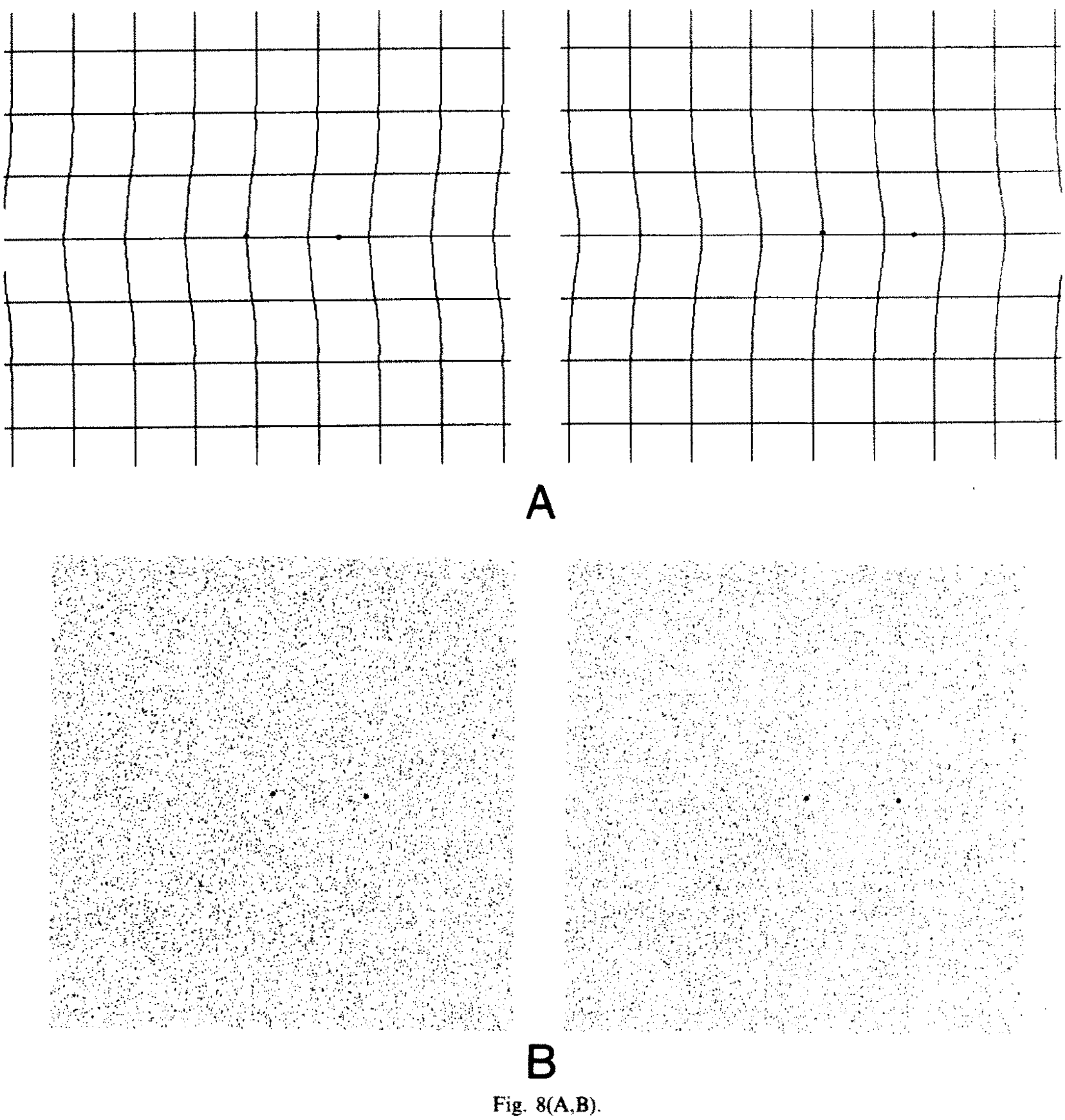
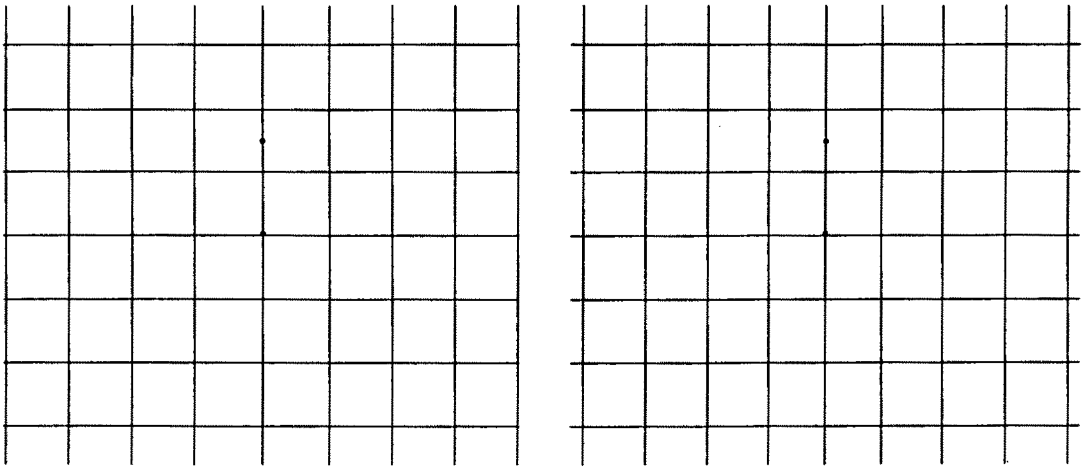
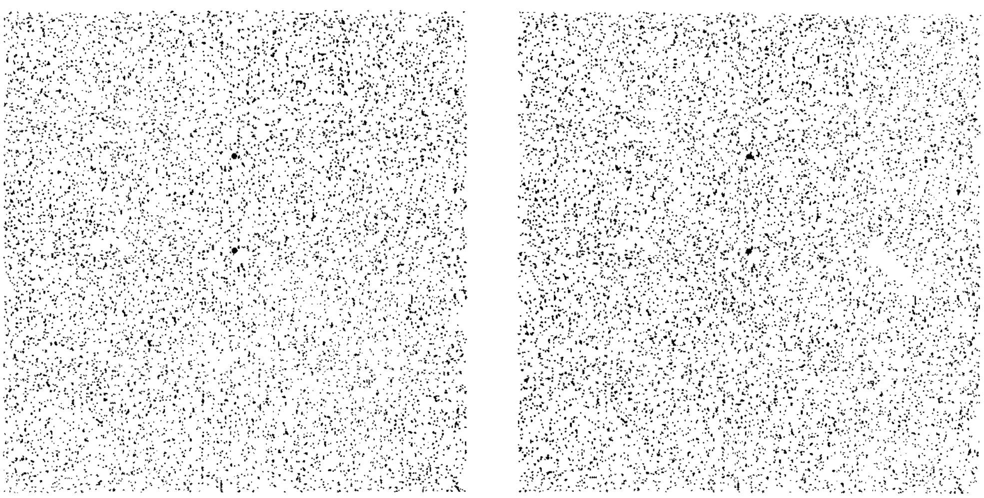


Fig. 8(A,B).



C



D

Fig. 8. Horizontal Gaussian ridges in (A) and (B); vertical Gaussian edges in (C) and (D).

not strictly determined by the disparity distribution.

Table 4 shows the responses for the grid stimuli. The values in boldface indicate depth judgments consistent with the relative stereo disparities. Consider the case of the slanted plane *h*, where disparity increased from south to north. The N probe location should have been seen as farther, but zero "farther than" judgments were in fact recorded, and likewise zero "nearer than" judgments for the corresponding S probe location.* Similar results were obtained for slanted plane *v*. It is remarkable that when the dots were embedded in a surface which had a constant gradient of disparity the apparent relative depth of the probe and reference dots collapsed. Points that were readily seen as lying at different depths when viewed in isolation appeared equidistant when embedded in the constant gradient, but seemingly unslanted, grid.

For the nonplanar cases, the edge and ridge, the data are in better accordance with the stereo information, and generally better for the *h* than the *v* surfaces. This anisotropy has been reported earlier by Tyler (1973), Wallach and

*Several of the relative depth responses were actually *opposite* that predicted by the stereo disparities. We conjecture that this was due to illusory linear perspective caused by stereo depth constancy compensation. While the grid lines were horizontal and vertical in each half-image, the fused grid appeared to be trapezoidal rather than rectangular, presumably because of apparent length was scaled with increasing disparity. The rectangular grid appeared distorted by linear perspective. The slant implied by the perspective, of course, was opposite that implied by the disparity gradient. This effect suggests to us that stereo size constancy operates independently of processes responsible for apparent depth.

Table 4. Percentage of responses that the probe point is nearer than (<), or equidistant (=), or farther than (>) the central reference point, as in Table 1. The probe and reference points were both embedded in a stereo surface, in this case rendered by a square grid (see Fig. 8). The relative depth judgment predicted by the relative stereo disparities is in bold

	Probe location											
Stereo surface	N			S			E			W		
and orientation	<	=	>	<	=	>	<	=	>	<	=	>
Slanted plane *h*	33	67	**0**	**0**	73	27	0	**100**	0	0	**100**	0
Gaussian edge *h*	0	0	**100**	**93**	0	7	0	**100**	0	0	**100**	0
Gaussian ridge *h*	0	0	**100**	0	0	**100**	0	**100**	0	0	**100**	0
Slanted plane *v*	27	**73**	0	0	**93**	7	0	80	**20**	**0**	100	0
Gaussian edge *v*	20	**80**	0	0	**87**	13	0	60	**40**	40	53	7
Gaussian ridge *v*	20	**80**	0	0	**93**	7	0	60	**40**	0	87	**13**

Bacon (1976), and Gillam *et al.* (1984) in depth detection tasks and by Rogers and Graham (1983) in the Craik–O'Brien–Cornsweet effect for stereopsis. Note that the depth of the Gaussian edge *v* was detected with only slightly better success than the slanted plane *v*.

We conclude that while depth can be encoded "directly" from disparity for isolated disparity points, *when those point are perceived as lying on a surface, their depth depends on the perceived depth of the surface*, which might happen to be negligible, either because it is a "featureless" field of stereo points in the absence of monocular 3D cues, or there are contradictory monocular cues.

The dramatic influence of the monocular grid is apparent in comparing the grid data in Table 4 with the corresponding random dot surface data in Table 5. The grid seemingly masked or "flattened" the depth undulation indicated by the disparity values. Significantly, the depth in the slanted plane stimuli, particularly in the *v* orientation, remained more difficult to detect than in the ridge and edge stimuli, even in the absence of a contradictory monocular 3D interpretation (of an unslanted rectangular grid). Ninio and Mizraji (1985) similarly observed that structured stereograms are less accurately perceived in 3D than unstructured (they used rectilinear grids as well). We interpret this as due to the conflicting monocular interpretation provided by the grids beyond the issue of the ineffectiveness of planar disparities.

GENERAL DISCUSSION

The 3D interpretation in these binocular stimuli was governed largely by the monocular cues. This is not to be construed as evidence of simple dominance of monocular over stereo cues, however. Instead, we believe that these planar stimuli happened to be particularly rich in monocular 3D cues, especially perspective and foreshortening, and particularly poor in stereo information due to our relative insensitivity to constant disparity gradients in the absence of disparity contrast. Stereo depth derives most effectively from disparity contrast; when disparity varies linearly it is dramatically less salient, despite large overall variations in disparity. In the absence of competing monocular cues a uniform gradient of disparity does effectively yield stereo depth, thus we do not conclude that stereopsis is wholly "blind" to constant disparity gradients. Rather, we suggest that depth interpretation from stereopsis is effectively reconciled with that from other sources primarily in terms of surface curvature and depth discontinuity features, and since our stimuli were devoid of these features, the monocular interpretation dominated.

Table 5. Relative depth judgments, as in Table 4, but for a surface depicted by a dense random dot pattern (see Fig. 8)

	Probe location											
Stereo surface	N			S			E			W		
and orientation	<	=	>	<	=	>	<	=	>	<	=	>
Slanted plane *h*	27	0	**73**	**73**	7	20	0	**73**	27	0	**67**	33
Gaussian edge *h*	7	13	**80**	**100**	0	0	0	**100**	0	0	**93**	7
Gaussian ridge *h*	0	0	**100**	0	0	**100**	0	**100**	0	0	**100**	0
Slanted plane *v*	40	**60**	0	0	**60**	40	0	7	**93**	**53**	7	40
Gaussian edge *v*	0	**100**	0	0	**100**	0	0	0	**100**	**100**	0	0
Gaussian ridge *v*	0	**100**	0	7	**93**	0	0	0	**100**	0	0	**100**

The fact that stereo depth must compete with monocular depth even in simple experimental stimuli likely accounts for several depth phenomena reported earlier. Westheimer (1979) and McKee (1983) observed that when two vertical lines, projected at different disparities, are connected by horizontal lines to form a square, the threshold for detection of the depth difference is greater than when only the two vertical lines are presented. McKee (1983) suggested that the effect was due to the lines being connected into a perceptual whole. Mitchison and Westheimer (1984), studying variations on this configuration, demonstrated that the detection thresholds were elevated most when the disparities varied linearly (according to a slanted plane). They use the term "salience" to refer to a local weighted sum of disparity first differences between a given point and its neighbors which scales roughly inversely with the separation of stereo features. [This notion quantifies Gogel and Mershon's (1977) "adjacency effect".] Accordingly, local variations in salience (i.e. second differences of disparity) would reveal deviations from planarity in the corresponding surface. A slanted plane would present points of equal salience, and consequently of zero apparent variation in depth. Gillam *et al.* (1984) observed, in these terms, that depth derives most readily from places of high "salience".

But Mitchison and Westheimer (1984) also said that more is involved in the perception of depth from disparity, since their proposal cannot account for the dramatic extinction of depth in the simple case of the slanted square compared to only the vertical lines of the square. McKee (1983) regarded this as a figural connectivity issue, recall. We believe McKee was close to the mark: it is not the connectivity *per se* that is important (as Mitchison and Westheimer demonstrated) but the fact that the connectivity helped induce a monocular figure, a square, that has a compelling 3D interpretation. The square suggested a plane of zero slant, which dictated that the two vertical sides of the plane are equidistant from the viewer. The following illustrates the dramatic influence a monocular interpretation has on the eventual depth percept.

An ellipse, seen from a particular viewpoint, foreshortens to a circle in orthographic projection—e.g. an ellipse of 2:1 aspect ratio rotated about its minor axis to a slant of 60°, so that the major axis foreshortens by a factor of 0.5 (the cosine of 60°). A 2:1 rectangle would likewise foreshorten to a square. The stereograms in Fig. 9 depict concentric ellipses (and rectangles) lying on a plane of 60° slant. A compelling monocular 3D interpretation would be of a tunnel or funnel extending in depth from periphery to center. Seven subjects, naive to the experimental design, interpreted the stereograms accordingly, with the innermost circle (or square) seen as further than the outermost. While some observers noted that the outermost circle (or square) appeared slightly slanted, the apparent slant vanished towards the innermost. Apparent depth increased radially towards the center of the pattern rather than from right to left, despite the fact that the vertical meridian was at zero disparity. When the subjects were subsequently told that the stimuli corresponded to foreshortened ellipses and rectangles lying on a slanted plane, some subjects could see the slanted plane, while curiously others could not.

Figure 10 is, we believe, a particularly effective demonstration of the monocular influence. The lines are coplanar, i.e. increase linearly in disparity from left to right. The 3D impression, however, is of a corridor extending in depth, bordered on either side by columns of vertical lines or stakes. In the apparatus the innermost lines on either side of the vertical meridian had stereo disparities of $\pm 11'$; the outermost lines had disparities of $\pm 51'$. It is remarkable that the line with $-11'$ disparity appeared more distant than the line of disparity $+51'$. This apparent disregard for stereo disparity is far more blatant than that reported by Mitchison and Westheimer (1984), where thresholds were elevated by only a few minutes of arc. The difference, we suggest, is that figure 10 offers a far more compelling monocular 3D interpretation. But it is also noteworthy that experienced stereo observers can also discern the true stereo depth of the component lines with scrutiny, especially in Fig. 10, as if the monocular depth interpretation can be selectively disregarded.

The final observation we offer concerns interactions between stereopsis and monocular interpretations in the case where the stereo disparities suggest a highly salient curvature feature. In Fig. 11 the monocular interpretation is of a slanted plane, but the stereo disparities correspond to a 2D Gaussian in depth protruding towards the viewer. Note that the disparities are symmetrically distributed over the two half-images so that the fused "cyclopean"

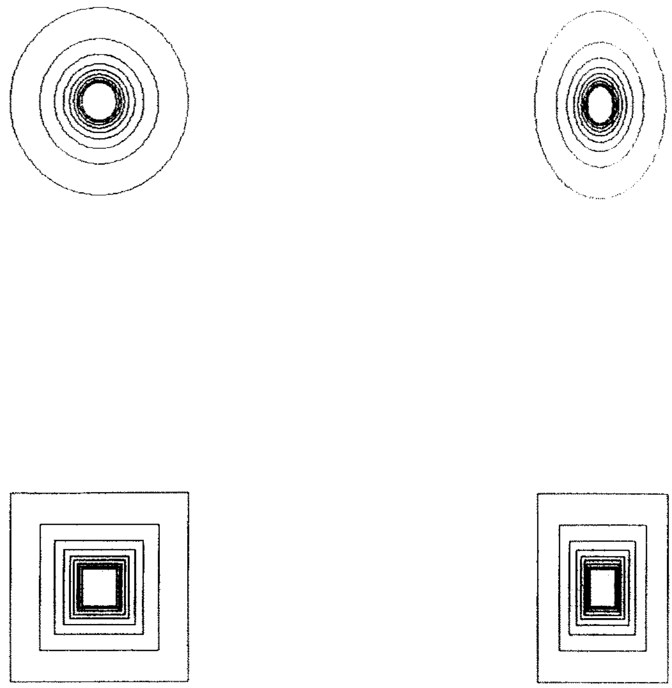


Fig. 9. Coplanar ellipses and rectangles, 2:1 aspect ratio and slanted 60°, in orthographic stereoscopic projection. A compelling monocular interpretation is of tunnels with circular and square cross-section seen in perspective.

Fig. 10. Lines on a common plane slanted 60°, but seen as a corridor in depth, as suggested monocularly.

image consists of straight lines, suggesting a slanted rectangular grid in perspective. We find that observers vary considerably in their interpretation of such a rivalrous figure, some seeing only a slanted plane, others seeing a plane at first then gradually becoming aware of a phantom protrusion in the center of the stereogram. Others achieve the nonplanar interpretation only after studying the random-dot stereogram version of the same Gaussian-shaped feature (Fig. 12) then re-examining the grid stereogram. Depth appears to be the end consequence of

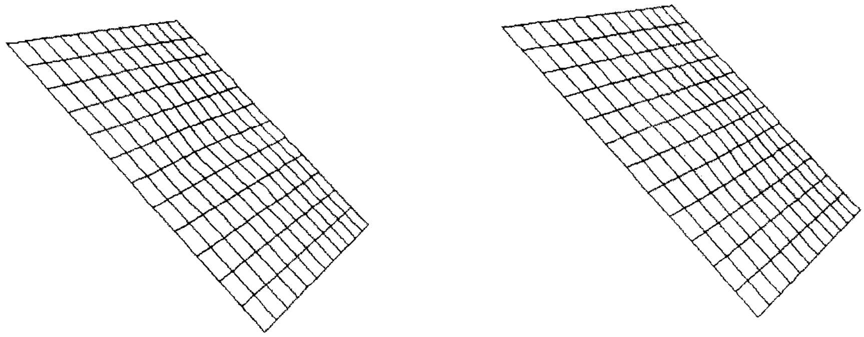


Fig. 11. A rivalrous pattern, monocularly a slanted plane, and stereoscopically a 2D Gaussian in depth.

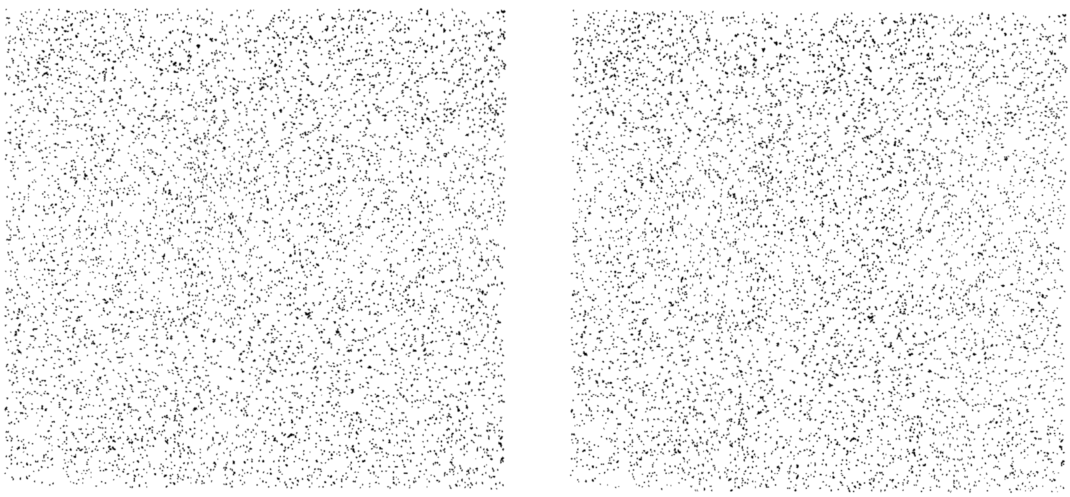


Fig. 12. The random-dot stereogram of the Gaussian in depth in Fig. 11.

a process that involves substantial "inference" or interpretation, that one sees depth according to the interpretation of 3D surface shape that one imposes. In that regard stereopsis is but one source of 3D shape information, and not necessarily the compelling one.

This series of experiments suggests that monocular cues have a stronger role in 3D perception than perhaps has been assumed. Likewise, stereopsis plays a much weaker role in the determination of depth across planar surfaces than expected. For very simple stereograms, an isolated pair of lines or points, say, the depth is indeed governed by the stereo disparities. But the contribution of stereopsis to the 3D percept changes dramatically as the stereogram is made more complex. With sufficient disparity evidence to suggest a continuous surface it is the *spatial distribution* of disparities, and not their individual magnitudes, that governs the apparent shape and depth. Specifically, the spatial distribution is analyzed to detect curvature and sharp discontinuities. Planar arrangements of disparity are in this regard featureless. This conclusion is close to that of (Gillam *et al.*, 1984) and (Mitchison and Westheimer, 1984) regarding the weak apparent depth associated with constant disparity gradients. In work reported elsewhere (Stevens and Brookes, 1988) we further conclude that surface curvature and discontinuity features are the primitive surface descriptors with which the visual system integrates stereo information with that contributed from monocular sources. In terms of spatial derivatives, we propose that the effective stereo features correspond to places where the second spatial derivatives are non-zero. The corollary is that neither the gradient (first spatial derivatives) nor the zeroeth derivatives (the raw disparity values themselves) are accessible as local surface shape descriptors. That is, neither slant nor relative depth is extracted directly from the disparity distribution across a surface. But, we must emphasize, relative depth is extracted from simple discontinuous configurations, such as between discrete, isolated items and across edges. And binocular vision undeniably provides absolute range information as well, particularly from convergence angle (Ritter, 1979) and in conjunction with motion parallax (Johansson, 1973). But we propose that range perception, which is most accurate in the near field (up to 2 m) under conditions of precise stereoscopic fixation, subserves motor functions such as locomotion and manipulation and *not* the perception of surface relief or 3D shape.

REFERENCES

Attneave F. and Frost R. (1969) The determination of perceived tridimensional orientation by minimum criteria. *Percept. Psychophys.* **6**, 391–396.

Gillam B. J. (1968) Perception of slant when perspective and stereopsis conflict: Experiments with aniseikonic lenses. *J. exp. Psychol.* **78**, 299–305.

Gillam B. J., Flagg T. and Finlay D. (1984) Evidence for disparity change as the primary stimulus for stereoscopic processing. *Percept. Psychophys.* **36**, 559–564.

Gogel W. C. and Mershon D. H. (1977) Local autonomy in visual space. *Scand. J. Psychol.* **18**, 237–250.

Gregory R. (1970) *The Intelligent Eye.* Weidenfeld & Nicolson, London.

Johansson G. (1973) Monocular movement parallax and near-space perception. *Perception* **2**, 135–146.

Julesz B. (1971) *Foundations of Cyclopean Perception.* Chicago Press, Chicago.

Longuet-Higgins H. C. (1982a) The role of the vertical dimension in stereoscopic vision. *Perception* **11**, 337–386.

Longuet-Higgins H. C. (1982b) Appendix to paper by John Mayhew entitled: "The interpretation of stereo-disparity information: the computation of surface orientation and depth". *Perception* **11**, 405–407.

Marr D. (1982) *Vision*. Freeman, San Francisco, CA.

Mayhew J. (1982) The interpretation of stereo-disparity information: the computation of surface orientation and depth. *Perception* **11**, 387–403.

McKee S. P. (1983) The spatial requirements for fine stereoacuity. *Vision Res.* **23**, 191–198.

Mitchison G. J. and Westheimer G. (1984) The perception of depth in simple figures. *Vision Res.* **24**, 1063–1073.

Morrison J. D. and Whiteside T. C. D. (1984) Binocular cues in the perception of distance of a point source of light. *Perception* **13**, 555–566.

Ninio J. and Mizraji E. (1985) Errors in the stereoscopic separation of surfaces represented with regular textures. *Perception* **14**, 315–328.

Ono H. and Comerford J. (1977) Stereoscopic depth constancy. In *Stability and Constancy in Visual Perception: Mechanisms and Processes*. (Edited by W. Epstein). Wiley, New York.

Perkins D. N. (1972) Visual discrimination between rectangular and nonrectangular parallelopipeds. *Percept. Psychophys.* **12**, 396–400.

Prazdny K. (1983) Stereoscopic matching, eye position, and absolute depth. *Perception* **12**, 151–160.

Ritter M. (1977) Effect of disparity and viewing distance on perceived depth. *Percept. Psychophys.* **22**, 400–407.

Ritter M. (1979) Perception of depth: processing of simple positional disparity as a function of viewing distance. *Percept. Psychophys.* **25**, 209–214.

Rogers B. J. and Graham M. E. (1983) Anisotropies in the perception of three-dimensional surfaces. *Science, N.Y.* **221**, 1409–1411.

Schriever W. (1925) Experimentelle Studien uber das stereoskopische Sehen. *Z. Psychol.* **96**, 113–170.

Shepard R. N. (1981) Psychophysical Complementarity. In *Perceptual Organization* (Edited by M. Kubovy and J. Pomerantz). Lawrence Erlbaum, Hillsdale, N.J.

Smith A. H. (1965) Interaction of form and exposure time in the perception of slant. *Percept. Motor Skills* **20**, 481–490.

Smith O. W. and Smith P. C. (1957) Interaction of the effects of cues involved in judgements of curvature. *Am. J. Psychol.* **70**, 361–375.

Smith P. C. and Smith O. W. (1961) Veridical perceptions of cylindricality: a problem of depth discrimination and object identification. *J. exp. Psychol.* **62**, 145–152.

Stevens K. A. (1981a) The visual interpretation of surface contours. *Artificial Intell.* **217**, 47–74.

Stevens K. A. (1981b) The information content of texture gradients. *Biol. Cybernet.* **42**, 95–105.

Stevens K. A. (1983a) Surface tilt (the direction of slant): a neglected psychophysical variable. *Percept. Psychophys.* **33**, 241–250.

Stevens K. A. (1983b) Slant-tilt: the visual encoding of surface orientation. *Biol. Cybernet.* **46**, 183–195.

Stevens K. A. (1984) On gradients and texture "gradients". Commentary on: Cutting & Millard 1984. Three gradients and the perception of flat and curved surfaces. *J. exp. Psychol. Gen.* **113**, 217–220.

Stevens K. A. and Brookes A. (1988) Theory of depth reconstruction in stereopsis. Submitted.

Tyler C. W. (1973) Stereoscopic vision: cortical limitations and a disparity scaling effect. *Science, N.Y.* **181**, 276–278.

Wallach H., Gillam B. and Cardillo L. (1979) Some consequences of stereoscopic depth constancy. *Percept. Psychophys.* **26**, 235–240.

Wallach H. and Bacon J. (1976) Two forms of retinal disparity. *Percept. Psychophys.* **19**, 375–382.

Wallach H. and Zuckerman C. (1963) The constancy of stereoscopic depth. *Am. J. Psychol.* **76**, 404–412.

Westheimer G. (1979) The spatial sense of the eye. *Invest. Ophthal. visual Sci.* **18**, 893–912.

Wheatstone C. (1852) On some remarkable, and hitherto unobserved phenomena of binocular vision. *Phil. Mag. Ser.* **4**, 504–523.

Yellott J. I. and Kaiwi J. L. (1979) Depth inversion despite stereopsis: the appearance of random-dot stereograms on surfaces seen in reverse disparity. *Perception* **8**, 135–142.

Youngs W. M. (1976) The influence of perspective and disparity cues on the perception of slant. *Vision Res.* **16**, 79–82.