CIS-TR-86-05

Integrating Stereopsis with Monocular Interpretations of Planar Surfaces

Kent A. Stevens
Allen Brookes

Department of Computer and Information Science
University of Oregon

INTEGRATING STEREOPSIS WITH MONOCULAR INTERPRETATIONS OF PLANAR SURFACES[1]

Kent A. Stevens and Allen Brookes
Department of Computer Science
University of Oregon
Eugene, OR 97403

Abstract. In a series of experiments in which stereo and monocular 3D interpretations were made mutually inconsistent, stereopsis was found to integrate with monocular depth only when disparities varied nonlinearly spatially. The apparent spatial disposition, including local surface orientation and depth ordering, was dominated by the monocular interpretation, even when stereo information provided strongly contradictory information. We conclude that stereo and monocular vision is integrated primarily on the basis of topographic features, not scalar depth values. Moreover, in interpreting depth from stereograms, we found a stereoscopic analogue to the familiar brightness contrast effect. The depth interpretation of stereo disparity information is roughly analogous to edge detection, with insensitivity to constant disparity gradients analogous to our insensitivity to linear (ramp-like) intensity distributions. This rules out several current models of spatial integration, and suggests new computational strategies.

1. Introduction

To what extent does stereo information constrain the perceived shape and spatial orientation of static surfaces? Under natural viewing circumstances, binocular observation of a near surface (*e.g.* less than two meters distant) seems to allow an observer to precisely localize the surface in depth. The perception of absolute distance from stereopsis is well calibrated in human vision out to at least 2 m (*e.g.* Wallach & Zuckerman, 1963; Ritter, 1977, 1979; Morrison & Whiteside, 1984), and in the near field the perception of depth intervals from disparity intervals are scaled appropriately according to distance (see review of stereo depth constancy in Ono & Comerford, 1977 and also Wallach, Gilliam & Cardillo, 1979). That is, at short range, both absolute depth and increments in depth are perceived with good accuracy. How then does near-field stereopsis integrate with monocular three-dimensional (3D) interpretations such as derived from figure 1? Given the sound geometrical basis for determining absolute and relative distances from stereo disparity, and the comparatively weak perceptual assumptions needed to infer a specific 3D interpretation from a monocular image, one would expect stereopsis to dominate over the less reliable monocular information. This study and others suggest to the contrary: monocular configurations may dominate the resulting 3D interpretation over stereopsis, even in the range where stereopsis has been demonstrated to be most accurate.

Usually, under experimental conditions, stereopsis allows for more veridical 3D interpretations over those derived from merely monocular information such as linear perspective, texture, and shading (*e.g.* Smith & Smith, 1957, 1961; Smith 1965). Contradictory results were reported by Youngs (1976), however, where stereo disparity had no significant effect on the the apparent slant of rectangles. Youngs (1976) questioned "why the disparity coding fails so miserably" in those experiments. The failure of stereo disparity to necessarily dictate apparent depth relationships in a binocular image has been known since Wheatstone (1852), and demonstrated *e.g.* in studies where reversed-disparity stereograms pit stereo disparity against monocular cues (Schriever, 1925; Gregory, 1970; Yellott & Kaiwi, 1979). Recently McKee (1983) and Mitchison and

---

[1] Supported by Office of Naval Research Contract N00014-K-84-0533.

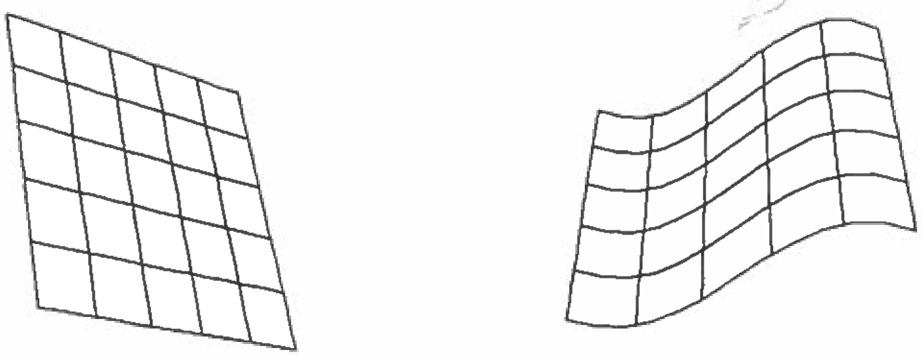

Figure 1. Monocular configurations that evoke definite 3D interpretations.

Westheimer (1984) report that the local arrangement or connectivity of lines in a simple stereogram influences apparent depth (see discussion later). One of the most dramatic demonstrations of the difference between disparity and depth is provided by Hochberg's random-dot stereogram depiction of a Necker cube (see Julesz, 1971, p. 163). Although the cube is at constant disparity against the background, it appears three-dimensional, that is, to fill a volume of space, and to reverse in depth. Clearly stereo disparity is not taken as the basis for the depth seen in the cube. While stereo disparity can induce an impression of depth, apparent depth is not functionally related to disparity according to any strict, geometrically-derived functional relationship, such as suggested recently (Mayhew, 1982; Longuet-Higgins, 1982a, b; Prazdny, 1983). Gillam, Flagg and Finlay (1984) have forcefully demonstrated this point, showing that apparent depth is evoked most effectively by sharp discontinuities in stereo disparity (see also Gogel, 1956, 1972; Gulick & Lawson, 1976).

But since stereopsis can provide a vivid impression of depth, particularly in simple line stimuli (such as used by McKee, 1983, and Mitchison and Westheimer, 1984) and those same line stimuli might also have a monocular interpretation in 3D, how are the two depth impressions integrated? That is, how does stereo and mono depth integrate? Gillam (1968) showed that when perspective and stereopsis are in opposition (in reversed-disparity stereograms) compromise slant judgments result, and quite recently Dosher, Sperling and Wurst (1986) have modelled similar monocular/binocular cue interactions by linear-weighted summation. We suspect that there are limits to the human visual system's ability to resolve the integration problem by linear combinations of cues, however. To begin, how does the visual system reconcile the situation posed by binocular observation of figure 1, above? Clearly the stereo information tells about the depth, curvature, and orientation of the sheet of paper on which the figure is printed; the stereo information is contradictory to the monocular 3D interpretation. Of course, when one examines the figure binocularly, one might discount the conflicting stereo disparities as being consistent with the surface of the printed page but not with the surface *depicted* by the figure, which appears to float indeterminately in depth roughly at the distance of the page, but somehow independent of it. The conflict, it might then be argued, is settled at a rather high level, where one chooses to see the printed figure "as if" it were in space. But this explanation is less plausible, however, under experimental conditions where the plane of projection is not visible. A Wheatstone-style stereo display presenting luminous points or lines against a totally dark background is particularly effective at generating a vivid sense of reality, since the observer sees only the binocular image in the empty volume of space immediately in front of the observer, without any other reference surfaces visible.

On what basis does the visual system integrate (*i.e.* combine, reconcile) the 3D information derived from stereopsis with that derived from other sources? Here the alternative sources we examine are static monocular cues to surface shape and orientation, but of course the question can be broadened to include, say, motion. Our experiments suggest that the visual system does not reconcile certain types of conflict between the 3D information implicit in the stereo disparities, and the 3D interpretation derived monocularly. The findings might suggest the rivalry between monocular and stereo interpretations are resolved in favor of the monocular, but we prefer to interpret them as evidence that stereopsis does not process certain types of disparity gradient information, and the monocular interpretation was taken in the absence of detected information to the contrary. In either case, the results argue against certain proposals for depth integration that otherwise seem intuitive, attractive, and computationally well founded. Returning to whether the integration can be modelled by linear-weighted summation or not, the question is, naturally, what is the nature of the evidence that is being linearly or nonlinearly weighted. Dosher *et al.* (1986) model the weighting of strengths of percept, which, while quite successful, leaves open what "strength of percept" might mean computationally. Does it mean, for instance, the apparent distribution of depth values consistent with each cue? That is, are depth values from different cues algebraically weighted and summated? Here we believe the answer is no. In the current study we found that depth (by which we mean specifically distance information) seems to play little direct role in the 3D integration process. That is, depth is almost an epiphenomenon of the process of 3D cue integration, a summary statement of how the surfaces seem to fit together from the evidence provided by the different cues (stereopsis, motion, monocular interpretations, and so forth). The integration might have a large component of linear-summation, but the question we addressed is what specifically is summated.

This study actually began with the interesting suggestion (W. Richards, personal communication, 1985) that the 3D interpretation of stereo disparity and monocular information might be mutually constraining, in a manner analogous to interactions between stereopsis and motion parallax (Johansson, 1973; Richards, 1985). Orthographic projections are ambiguous in depth; stereopsis might disambiguate which points are near and which are far. It could also, in principle, provide 3D information that could verify a given monocular interpretation based on assumptions. For example, there is an apparent tendency to interpret an oblique intersection in a monocular image as corresponding to a right-angle intersection in 3D (see discussion below). Although assuming perpendicularity allows one to infer the orientation of the corresponding 3D intersection, it is not "reflective" (Ullman, 1979), *i.e.* not directly verifiable in the monocular image. Stereopsis could play a role in verifying whether the intersection is indeed perpendicular (see appendix for formulae showing how stereo disparity could be used to compute the actual 3D angle of intersection). Likewise, it is conceivable that monocularly-derived 3D information could constrain the interpretation of stereo disparities, particularly the relationship between disparity and relative distance. We therefore set out to examine the actual influence of stereopsis on monocular 3D interpretations and vice versa. We started with very simple configurations, and found unexpected interactions that had little apparent relationship to the notion of mutual constraint.

2. Experiments

2.1 Experiment 1: The Perpendicularity Constraint versus Stereopsis

It is well known that observers tend to interpret monocular images of intersecting lines or corners in 2D as right-angle intersections or corners in 3D (Attneave & Frost, 1969; Perkins, 1972; Shepard, 1981; Stevens 1983). In an experiment related to the current one, Stevens (1983) measured the apparent tilt (the direction of slant) of monocularly-presented crosses. The stimuli were merely intersecting pairs of straight line segments, where the ratio of line lengths and the angle of intersection were varied experimentally. A line segment with one end fixed at the point of intersection, the other free to rotate in the image plane could then be briefly superimposed on the cross, and subjects adjusted its orientation so that it coincided with the direction to which the normal would project in the image plane. The resulting tilt data corresponded closely to that predicted by assuming the intersection was a right angle in 3D and that the two line lengths, generally unequal in the image, were equal in 3D (Stevens, 1983). That experiment provided a baseline for judgments of the apparent surface tilt based on what we will refer to as the perpendicularity constraint. (Note that the 3D interpretation is apparently constrained by effectively presuming perpendicularity, but we make no suggestions of how that constraint is actually embodied in a perceptual computation.)

We then presented cross and grid stimuli in accurate stereo projection to examine possible mono/stereo interactions. When viewing a stereogram of a right angle intersection, the monocular 3D interpretation is consistent and redundant with that provided by the stereo disparity information. But if the stereogram presents an intersection that is *not* perpendicular, the monocular 3D interpretation would be inconsistent with that provided by the stereo disparities. Which spatial interpretation prevails, the monocular or the stereo, or some compromise?

Method

Apparatus: The stereo pairs were presented by a Wheatstone-style stereoscope using a pair of optically flat front-surfaced mirrors and two Tektronix 634 monochrome displays (which have flat 9x12 cm screens, 1100 line resolution, and less than 0.5% geometric distortion). The two monitor screens were perpendicular to their respective incident optic axes, positioned 38 cm from the observer along the optic axis and viewed with the corresponding angle of convergence of 9.8° that is correct for a 65 mm interpupillary separation (so that accommodation and vergence were consistent). Observers viewed through a pair of circular apertures which allowed a 6.4° radius field of view. The stimuli consisted of luminous lines against a dark background. The stereoscope was shielded from ambient light, and observed with normal interior light adaptation. The stereograms were generated dynamically by a Symbolics 3670 Lisp Machine; the monochrome monitors projecting the left and right images were driven independently by the red and green channels of a Robotic Systems, Inc. color frame buffer.

To generate a stereo pair, 2D projections were computed from left and right vantage points differing by the 9.8° convergence angle. The two images could be generated in either perspective or orthographic projection. In the perspective case the surface was computed as if physically situated 38 cm from the viewer; for the orthographic case 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. Note that while the perspective case was geometrically accurate, the orthographic stereo projection presents disparities consistent with the 9.8° convergence angle (appropriate for the 38 cm viewing distance) and yet images consistent with very distant observation.

Stimuli: The stimuli consisted of stereo projections of intersecting line segments, where the angle of intersection was 90° plus a skew angle of 0, 15, 30, or 45° (*i.e.* 4 conditions). The experiment was to place in conflict the monocular interpretation of the intersection as perpendicular with disparity information that indicated otherwise. Two types of stimuli were shown: a simple cross consisting of two equal-length lines, 5 cm long, that met at their midpoints, or a 4x4 grid composed of equal-length line segments, each 5 cm long, intersecting at the given skew angle. Figure 2 shows cross and grid stimuli for the 0° skew case, and figure 3 stimuli of the same spatial orientation but 45° skew. Note that the grids were square when the skew angle was 0° and parallelograms otherwise.

The subjects uniformly fused the stimuli immediately, and reported seeing a planar surface in 3D; the binocular image was clear, sharp, and stable over eye movements. The left and right

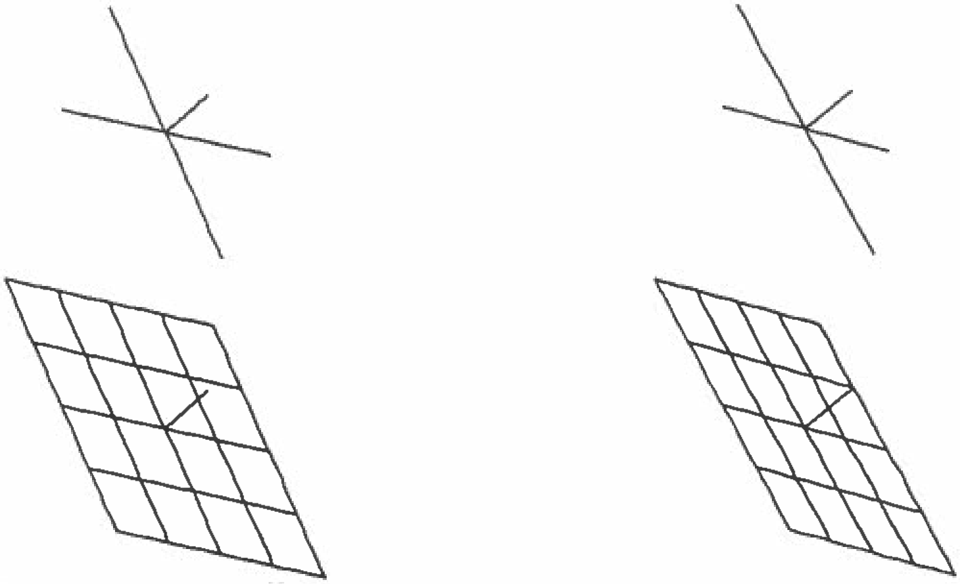

Figure 2. Examples of cross and grid stimuli, each with 0° skew angles. Note that the normal appears to project perpendicularly to the plane defined by the cross or grid.

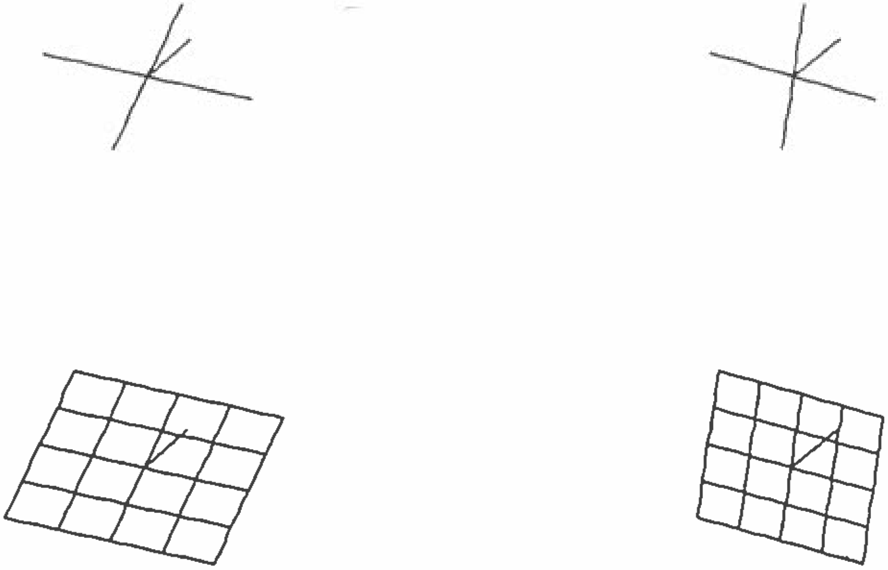

Figure 3. Cross and grid stimuli, with identical spatial orientation as in figure 2, but at 45° skew. Note that the "normals" do not appear perpendicular to the plane of the grid or cross.

images of the stereo pair were generated in orthographic projection. The slant (the angle between the normal to the plane and the line of sight) was held constant at 65°. The surface tilt (the direction of the normal measured in the image plane) was varied over three values to provide stimuli that were geometrically similar but superficially different. Finally, the cross or grid figures were presented at two angular orientations relative to the surface normal, as if the figures were rotated in 3D about their surface normal, again presenting different image projections. The observer thus saw intersecting lines in 3D from different viewpoints. The intention was to collapse the data across the three tilts and two rotations about the surface normal to provide six data values for a given condition of skew angle and stimulus type (*e.g.* Stevens, 1983). In an earlier pilot experiment, to meet concern that the continuous lines might be less adequate stereo stimuli, 10 subjects were presented with similar stimuli, where the lines were either dotted versus continuous. No significant differences were found between the dotted and continuous cases, therefore we used only continuous line drawings in this experiment. The experiment thus consisted of a sequence of 48 conditions: grid and cross stimuli at 4 skew angles, 3 tilts, and 2 rotations about the surface normal.

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 intersections in 3D would sometimes be right angles and at other times skewed. It was stressed that the projections would always appear foreshortened, and that they were to judge the actual angle of intersection in 3D. Their perpendicularity (P) judgment was indicated by depressing a mouse button, responding positive if the lines appeared within approximately 5° of perpendicular. After recording the response, a third line segment, computed to be the normal to the plane containing the two lines, was also projected stereoscopically. The subject was to then respond, by mouse button, whether that line appeared three-dimensionally perpendicular to the surface suggested by the grid or intersection (with the same criterion of roughly 5°). Unlimited presentation time was allowed, but subjects were urged to respond as quickly as they were able to reliably, particularly for the normal (N) judgment, since there was concern that given sufficient observation time one might eventually see the normal as being correctly oriented even though it did not appear so initially.

Results

The data across subjects for both the grid and intersection (both P and N judgments) showed no significant difference as a function of tilt, holding constant the skew angle and the rotation angle about the normal. The judgments were similarly independent of rotation angle; the means differed significantly only at one instance at the 0.05 level and one other at the 0.1 but not the 0.05 level. These two cases occurred for stimuli differing only in tilt, but the third tilt condition showed no significant difference.

With the data collapsed across tilt and rotation angle, the P and N judgments could be examined as a function of skew angle (table 1).

Table 1. Judgments of Perpendicularity (P) and Normal (N)

	ΣP	ΣN	$P\&N$	$P\&\neg N$	$\neg P\&N$	$\neg P\&\neg N$
SKEW-ANGLE = 0°						
Cross:	50	57	47	3	10	0
Grid:	54	49	44	10	5	1
SKEW-ANGLE = 15°						
Cross:	42	52	40	2	12	6
Grid:	54	50	46	8	4	2
SKEW-ANGLE = 30°						
Cross:	24	48	17	7	31	5
Grid:	43	37	25	18	12	5
SKEW-ANGLE = 45°						
Cross:	12	42	7	5	35	13
Grid:	41	21	11	30	10	9

In table 1, P refers to judgments in which the intersections appeared perpendicular (and $\neg P$ to cases where the intersections appeared skewed), and N refers to judgments in which the normals appeared correctly orthogonal to the plane of the intersection (and $\neg N$ to cases where the normal appeared incorrect). The totals for P and N are given in columns ΣP and ΣN; the other columns break down the totals into the four possible combinations of judgments of N and P. The column for $P\&\neg N$, for example, displays the totals across 10 subjects of trials where the intersections appeared perpendicular but the normals seemed incorrect.

Examining the ΣP column, note that with increasing skew angle the crosses lose their apparent perpendicularity (decreasing from 50 out of 60 trials for the true 90° case to only 12/60 trials when actually 45°) but the grid was still seen as perpendicular in 41/60 trials when the angle of intersection was actually 45°. Examining the ΣN column, the opposite trend was found: the number of trials in which the normal appeared correct diminished from 49/60 to 21/60 for the grid stimuli while remaining 42/60 for the cross. The ΣP responses for cross and grid stimuli were not significantly different for skew angles of 0° ($t(18) = 0.88$, $p > 0.2$) or 15° ($t(18) = 1.9$, $p < 0.1$), but were significant for skew angles of 30° ($t(18) = 2.36$, $p < 0.05$) and 45° ($t(18) = 3.66$, $p < 0.02$). Similar results were found for the mean N responses. For the cross stimuli it was found that stimuli differing by only 15° in skew angle barely reached significance, while those differing by 30° were reliably different (*e.g.* for skew-angles 15 versus 45°, $t(18) = 3.58$, $p < .01$). In contrast, for the grid data, the judgments varied much less as a function of skew angle, reaching significance only for skew angles of 0 versus 45° ($t(18) = 2.28$, $p < 0.05$) and for 15 versus 45° grid (t(18) = 2.28, $p < 0.05$).

Discussion

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 determined 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 monocular interpretation had determined the judgments, the intersections would have always appeared perpendicular and the normal would have appeared incorrect for all but the 90° case.

While the data fell somewhere between these two hypotheses, the monocular perpendicularity interpretation was markedly influential over the geometrically-correct stereo information. There was an apparently stronger tendency to "ignore" the stereo information in the grid stimuli compared to the cross stimuli, since even the 45° angle of intersection was judged to be a right angle 41/60 times (compared to the baseline of 54/60 for the true 90° angle of intersection). By contrast, the strong extinction of the P interpretation in the 45° cross (12/60) suggests that there was sufficient stereo information to reveal the actual three-dimensional angle of intersection. This difference in performance is most apparent in comparing the P&¬N and ¬P&N columns for the cross versus grid at 45°. The angle of intersection appeared skewed and the normal correct (¬P&N) in 35/60 trials for the cross and 10/60 for the grid. The intersection appeared perpendicular and the normal incorrect (P&¬N) in 4/60 trials for the cross and in 30/60 for the grid.

Figure 4 shows, for the grid stimuli, how increasing the skew angle increasingly distorts the monocular configuration and changes the apparent orientation of the plane of the grid in 3D. The normal is geometrically correct for 0° skew angle, and appears increasingly incorrect with increasing skew angle. (This figure was generated by producing a monocular projection of the corresponding stereogram stimulus.) The normal in figure 5a is correct for a 45° skewed grid, but

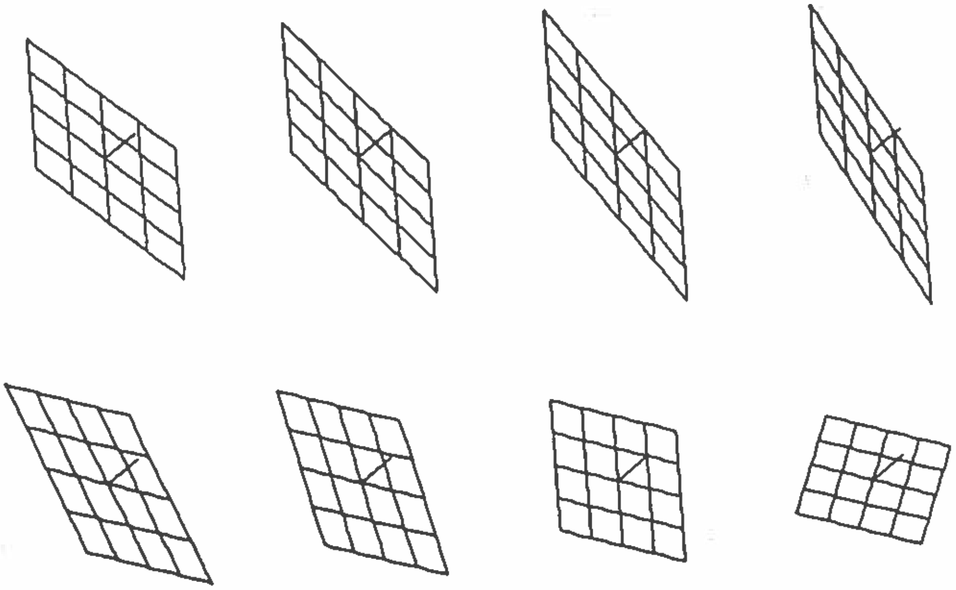

Figure 4. Family of grid stimuli of varying skew (from left to right: 0, 15, 30, and 45°) and rotation of the grid about the surface normal differing by 60 degrees in 3D.

clearly appears incorrectly oriented relative to the plane of the grid. In figure 5b, the line segment is more plausibly the image of the surface normal to the grid. It was computed by analyzing the 2D configuration as if it were the projection of a square grid (see Stevens 1983c, appendix, for formula). Figure 5b represents a close approximation to the monocularly perceived surface tilt and illustrates the difference between what is geometrically correct for a 45° intersection, and what one would perceive if that intersection were assumed to be 90°. In this case the difference in tilt amounts to 64° between the true (binocular) tilt and the monocular tilt. Likewise, the slant component of the surface orientation (the angle between the surface normal and the line of sight) is also influenced by assuming the intersection is 90°. All the stimuli were presented stereoscopically at 65° slant angle. But the skewed stimuli acquired different apparent slants. For example, the slant of the grid in figure 3 appears much less than 65°. The computed monocular slant for this stimulus, assuming it corresponds to a square grid, is only 38.5°, which is in good correspondence with the apparent slant seen in the binocular stimuli.

One might suggest that the P judgments in the case of the grid stimuli, which reflect a monocular interpretation of the intersections as perpendicular, do not necessarily arise from subjects actual judging perceived angles in 3D, but instead, positive responses to figures that resemble familiar grid patterns. And subjects sometimes remarked that they could, with more or less effort, interpret virtually any of the stimuli as involving right angles in 3D. Thus the P data might have been corrupted by superficial responses. On the other hand, we found in pilot experiments that subjects dramatically underestimated the magnitude of skew in the grid case. Two non-naive observers, presented with cross stimuli of unknown skew angle could accurately estimate the intersection angle to within 5° or so, and yet, for grid stimuli, repeatedly judged a 45° intersection as a right angle skewed only 15° or so. Furthermore, the diminished N judgements in the cross versus the grid stimuli are hard to reconcile by an "expectation for right angles" explanation, since it is one matter to see the intersections as perpendicular or not, and another to infer the spatial orientation consistent with perpendicularity. The finding was that skew angle

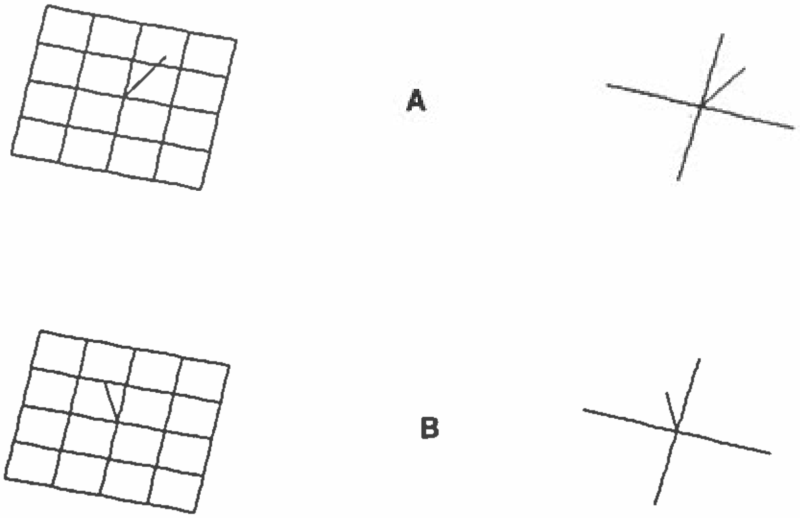

Figure 5. In *a* the normal is correct for the monocular projection of a cross skewed 45°. In *b* the normal is correct for the monocular projection of a right angle intersection.

had a large influence on the spatial orientation of the grid, which caused the binocular normal to appeared increasingly incorrect as skew angle increased. Given that the normal was an accurate stereo projection of the normal to the plane containing the lines regardless of their skew angle, we conclude that the monocular 3D interpretation was responsible.

How do the grid and cross stimuli differ? Geometrically, the grid configuration subsumes the cross: there are 8 lines versus 2, 16 intersection points rather than 1, but identical angles, lengths and so forth. Monocularly, therefore, the geometry of the foreshorted grid is redundant relative to that of the cross. But a grid also presents a regular arrangement of parallel lines, perhaps more compellingly suggestive of a surface than a mere intersection of two line segments. Regarding stereopsis, one might expect that the denser two-dimensional region of stereo disparity values provided by the grid would lead to a more compelling stereo interpretation. The fact that the grid data conformed less with the stereo and more with the monocular 3D interpretation was therefore unexpected.

Subjectively, the cross stimuli resembled a pair of vectors in space. The P judgment amounted to measuring the included angle and the N judgment to determining whether the normal, a third space vector, was mutually perpendicular to the other two vectors. The P judgment was apparently a focal, conscious task. Several subjects in experiment 1 remarked that they made the P judgment by mentally rotating the plane containing the two lines in order to "look straight down" on the intersection. Others imagined the two lines were intersecting runways or roads seen from the air. Eye movements along the two lines seemed helpful in judging their spatial orientation, particularly the N judgments. The grid stimuli tended to assume a monocularly-imposed spatial orientation appropriate for the assumption that the intersections were perpendicular, and this influenced the way the subjects saw the intersection angle and the normal to the grid. But if the visual system were to determine slant from stereo disparity, the grid stimuli, in providing a more extensive 2D region of disparity values, should have resulted in more accurate P and N judgments.

2.2 Experiment 2: Monocular Influence over Stereo Slant and Shape Constancy

This experiment used binocular stimuli in which the monocular and stereo information was contradictory. Consider an ellipse inscribed on a plane. Projected orthographically from a particular viewpoint, the ellipse foreshortens to a circle in the image. For example, an ellipse of 2:1 aspect ratio foreshortens to a circle from an oblique viewpoint such that the tilt aligns with the major axis and the slant is 60°, which foreshortens the major axis by a factor of 0.5 (the cosine of 60°). Likewise, a rectangle of 2:1 aspect ratio can be foreshortened to a square in orthographic projection. A circle or square is usually interpreted as lying in the image plane. The stereo disparities, however, correspond to a plane slanted 60° to the line of sight, hence the figure should appear slanted and elongated in depth. This is, of course, a matter of shape constancy; we examined the possible effect of the monocular interpretation on shape constancy.

Method

Stimuli: Stereo pairs were generated of elongated ellipses and rectangles projected from a perspective that foreshorted them to circles and squares[2]. As in experiment 1, the stereo projection corresponded to a 38 cm viewing distance and 9.8° vergence angle. The half-images were computed in orthographic projection. Holding slant constant at 60°, the tilt of the plane containing the slanted stimulus shape was either horizontal or vertical. The ellipses and rectangles were 1.7 by 3.4 cm (subtending 2.5 by 5°) for a 2:1 aspect ratio, centered in the binocular field of view. The stereo disparity was zero at the center and ±19′ at the far and near extremes of the stimulus

---

[2] Here we refer to the fused binocular image as a 2D projection in the sense of Julesz's (1971) "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.

Figure 6. Stereograms of an ellipse and a rectangle, each of 2:1 aspect ratio slanted 60° to the line of sight, horizontal tilt.

Figure 7. Same as figure 6 but with vertical surface tilt.

surface. The slanted ellipse (rectangle) stimulus could be compared with the corresponding circle (square), projected at zero stereo disparity and subtending the same size. Example stimuli, for 2:1 aspect ratio and horizontal surface tilt are shown in figure 6; figure 7 shows corresponding stimuli for vertical surface tilt.

A subsequent stimulus pattern was a family of concentric ellipses (rectangles) lying on a common plane of 60° slant (figure 8). A natural "monocular" interpretation of the apparent perspective in this figure is of a tunnel or funnel extending in depth from periphery to center. In figure 9 the surface tilt is vertical rather than horizontal. As a control, the minimal stereo stimuli for perceiving slant and elongation in depth was provided by projecting only the four extremal points of the ellipse. Geometrically, the left and right points were separated in 3D by twice the distance between the upper and lower points. A 60° slant angle provided a large gradient of stereo disparities; subjects had no difficulty achieving and maintaining stereo fusion of the stimuli.

Procedure: Seven subjects were used; four were naive, the other three were aware of the stimulus design and of the purposes of the experiment. The surface tilt was horizontal in the first series of presentations then vertical. In both series, subjects first compared the slanted ellipse with an unslanted circle at zero disparity. The aspect ratio of the ellipse was reduced from 2:1 until the ellipse appeared circular, using the circle as a reference. Unlimited observation time was provided, and free eye movements were allowed. The mean aspect ratio determined from the these presentations was then used in subsequent presentations of the four extreme points of a single slanted ellipse and the concentric slanted ellipses. Each stimulus figure was observed for several seconds then described verbally. After completing the ellipse stimuli with the underlying plane tilted both horizontally and vertically, the presentations were repeated with stimuli composed of rectangles instead of ellipses.

Results

The ellipse of 2:1 aspect ratio, horizontal tilt and 60° slant, had modest (10-30°) apparent slant and only slight elongation in depth. It was possible to make the ellipse appear satisfactorily circular when the aspect ratio was between 1.7:1 and 1.75:1 (see figures 8-10). These values are imprecise because the apparent slant and shape were variable. The stimulus would gradually, over 3-5 seconds, increase in apparent slant and simultaneously elongate in depth. It also appeared warped or distorted in depth to some observers.

The stereogram consisting of the four extrema points was seen accurately in 3D by all observers. The imaginary line connecting the left and right points appeared strongly slanted and longer than the vertical line connecting the upper and lower points, all in accordance with the stereo information. The family of coplanar, concentric, slanted ellipses, on the other hand, appeared to all subjects in 3D as a tunnel or funnel extending in depth, with the innermost circle as farther than the outermost. While the outermost circle appeared slightly slanted, the apparent slant vanished towards the innermost. Some observers interpreted the stereogram more as a funnel or cone than a tunnel, since the diameter appeared to diminish as the distance increased towards the center of the figure. Despite the fact that the vertical meridian was at zero disparity (with stereo disparities increasing to the left and decreasing to the right) apparent depth increased radially towards the center of the pattern rather than from right to left. All subjects saw the center as farther.

Different results were obtained when the slanted ellipse was oriented with vertical tilt, so that depth increased upwards. The ellipse of 2:1 aspect ratio, which foreshortens in the image plane to a circle, appeared in depth veridically as an ellipse of approximately 2:1 aspect ratio, and for most subjects no aspect ratio greater than 1:1 looked circular. The shape constancy for this orientation of the stereo disparity gradient was excellent. The slanted concentric ellipse stimuli generally suggested a tunnel, but some subjects interpreted the stereogram literally as concentric, slanted, ellipses all lying on a common plane.

Similar results were found with the rectangle stimuli. The horizontally-tilted slanted isolated rectangle was roughly square for an aspect ratio of approximately 1.7:1, and a vertically-

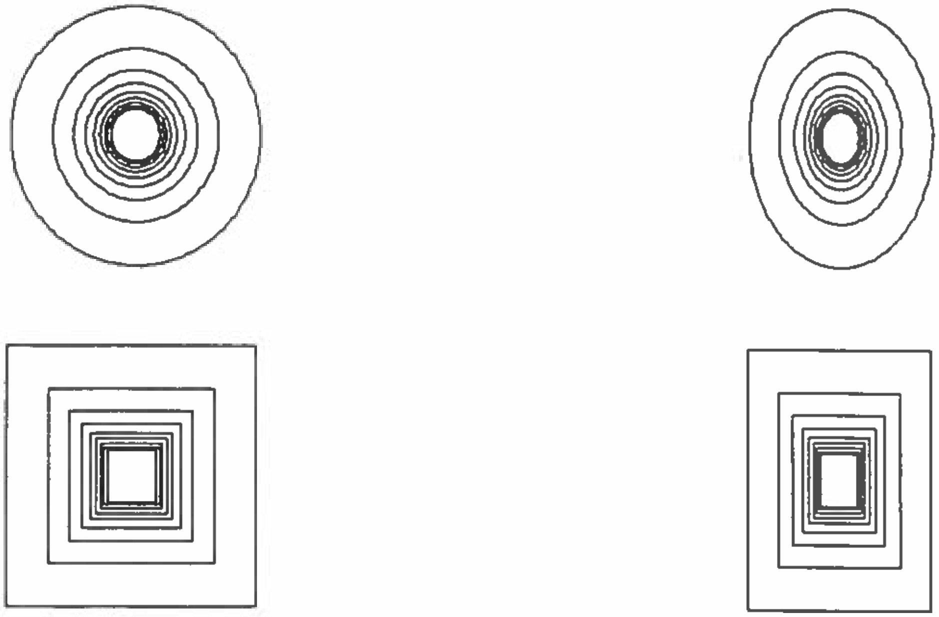

Figure 8. Concentric ellipses and rectangles of 1.7:1 aspect ratio slanted 60° horizontally.

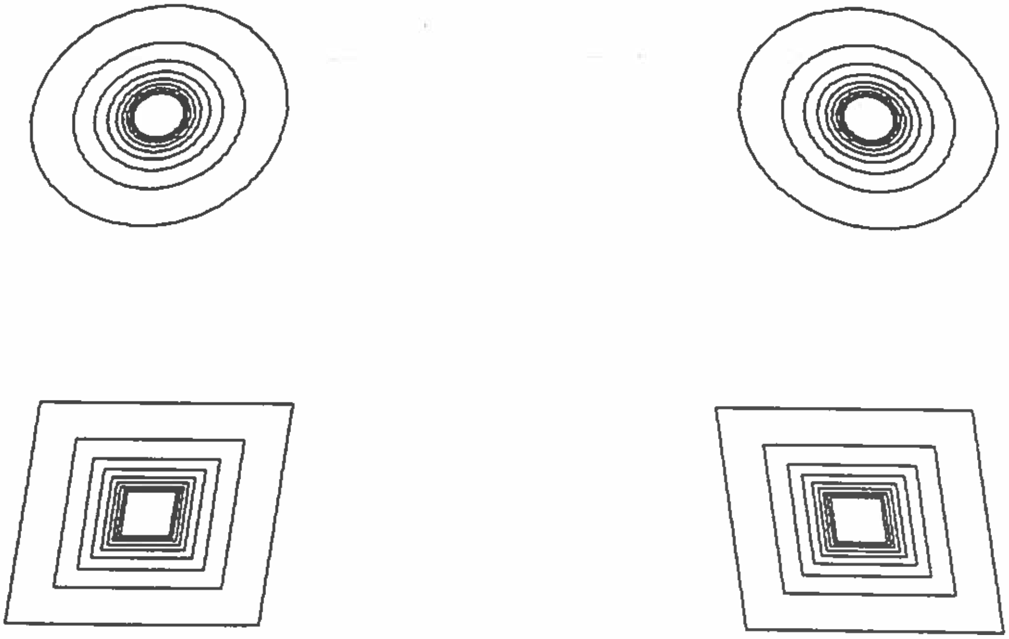

Figure 9. Same as figure 8 but with vertical surface tilt.

tilted slanted rectangle had a greater tendency to be seen as slanted and elongated, and often did not appear square until the aspect ratio was reduced to 1:1 in 3D. The concentric slanted rectangle stimuli was seen as a square tunnel or corridor, however. The outermost rectangle appeared slanted but the innermost appeared unslanted and approximately square when the aspect ratio was 1.5:1 (after a few seconds of observation, during which time the rectangle elongated in depth). The subjects were then told that the stimuli corresponded to foreshortened ellipses and rectangles lying on a slanted plane, whereupon some subjects saw the slanted plane, curiously others could not. Greater success was had in achieving the slanted-plane interpretation in the vertical case where depth increased upwards.

In light of these results and the earlier results in experiment 1, we had concern that somehow our stereo apparatus was not permitting sufficiently accurate stereopsis, so we implemented a stereo display of an isolated short line segment, floating like a needle against a black background, that could be rotated by the observer into any spatial orientation. The line segment had one endpoint fixed at zero disparity in the center of the field of view; the other endpoint could be adjusted in slant and tilt with 2.5° increments. The needle was immediately and dramatically three-dimensional, resembling a vector in space. Small changes in slant, *e.g.* of 65° and 67.5° were clearly distinct. The stereo apparatus was evidently able to support sufficiently precise spatial judgments for our purposes.

Discussion

The disparities in the stereograms varied continuously and linearly across the given stimulus surface, corresponding to a slanted plane viewed from only 38 cm. There was sufficient stereo information available to determine the spatial orientation of the stimulus plane, since the stereograms consisting of merely four points were immediately and accurately seen as lying on a slanted plane. But when the stereogram had an alternative monocular interpretation, that interpretation

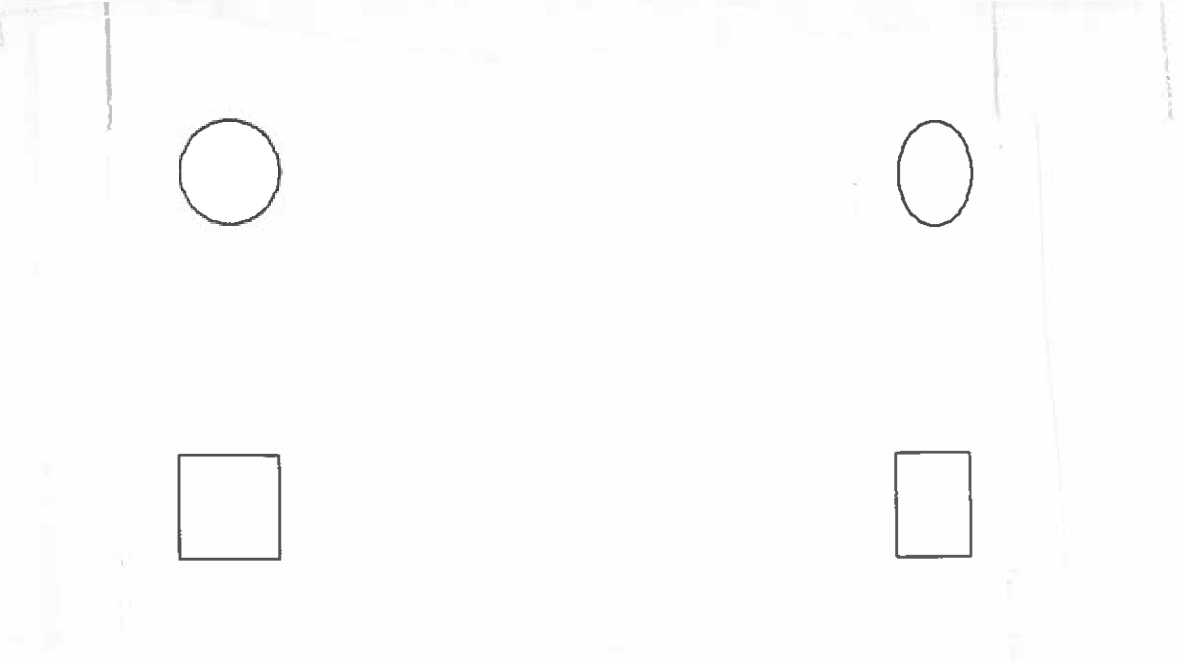

Figure 10. Same as figure 6, but with aspect ratio of 1.7:1 that makes figures appear roughly circular and square.

generally prevailed — as in experiment 1.

In a particularly clear demonstration of this pheomenon, a variation on the rectangle stereogram display was presented. In the first presentation only the vertical lines of the slanted rectangle was shown (figure 11a). The two lines appear at dramatically different depths, as one would expect on the basis of their stereo disparities. In the second presentation the vertical lines of the concentric rectangles were added (figure 11b). Recall that the rectangles are nested but all lie on a common plane that is slanted 60° to the viewer. The lines were thus coplanar and linearly ordered in depth from left to right according to their stereo disparities. The three-dimensional impression, however, was of a corridor extending in depth, bordered on either side by columns of vertical lines or stakes. The vertical meridian of the field of view corresponded to zero disparity and the innermost lines had stereo disparities of ±11′; the outermost lines had disparities of ±51′. It is remarkable that the line with -11′ disparity appeared more distant than the line of disparity +51′. The depth ordering, obvious in an isolated pair of lines, was very difficult to determine when additional lines were introduced lying on the same plane.

In both experiments 1 and 2 the distribution of stereo disparities was linear across the stimulus surface. The surface was viewed against a featureless black background through circular apertures set near the eyes, hence there were no step or edge discontinuities in stereo disparity. But the stimulus surfaces subtended several degrees of the field of view, over which disparities typically varied between ±20′, providing a significant constant gradient of stereo disparities. The absence of any background features is critical to this observation. The visual system was presented with stereo disparities that varied continuously along the lines of the stimuli, and unless one were to compare the disparities between specific points, such as the endpoints of the lines comprising the cross stimuli in experiment 1, or between pairs of dots in experiment 2, no "disparity contrast" was present. Recall that Youngs (1976) found that stereo disparity was ineffective in slanted-plane stimuli, and yet many earlier studies found that stereo disparity enhanced apparent three-dimensionality. Youngs' stimuli, however, had continuous and constant

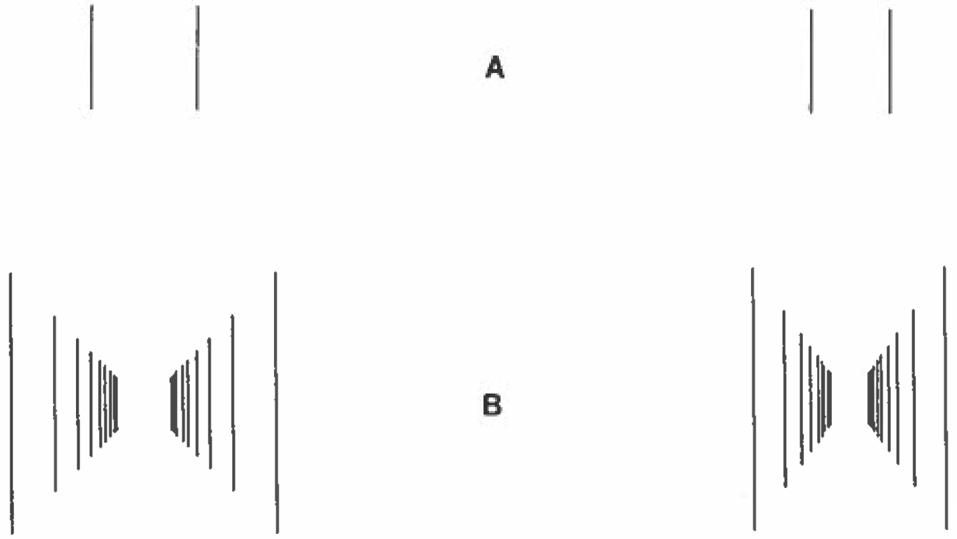

Figure 11. The depth ordering in *a* is apparently disrupted by the alternative monocular interpretation of a corridor when additional lines are added in *b*.

gradients of disparity while the stimuli in the other studies exhibited discontinuities or nonlinearities induced by surface curvature. Gillam *et al.* (1984) have concluded that depth derives most effectively from disparity discontinuities, and cite the stereo depth version of the Craik-O'Brien-Cornsweet illusion (Anstis, Howard & Rogers, 1978; Rogers & Graham, 1983) as supporting their conclusion.

Studying the 3D integration of mono and stereo is therefore confounded by the relative insensitivity of the visual system to stereo slant (*i.e.* linear disparity distributions that correspond to slanted planes). Our stimuli in experiment 2, where stereo information suggested a slanted plane but mono information suggested a tunnel or corridor, might well have been dominated by the mono interpretation simply because stereopsis was providing no information to the contrary. But this is likely not the whole story, because the visual system is not entirely insensitive to linear disparities. The curious phenomena reported by Westheimer (1979) and McKee (1983) where the connectivity of lines in a simple figure influences the stereo depth likely has two interwoven and confounding contributions. The basic finding was that when two vertical lines, projected at different disparities, were connected by horizontal lines to form a square, the threshold for detection of the depth difference was greater than when only the two vertical lines were presented. McKee (1983) suggested that the effect was due to the lines being connected by parts of a figure. Mitcheson and Westheimer (1984) studied variations on this configuration, and demonstrated that the detection thresholds were elevated most when the disparities varied linearly (according to a slanted plane). They concluded that perceived depth derives from spatially-weighted local differences in disparity, and proposed a computation for determining an "adjacency effect" (Gogel and Mershon, 1977) that scales roughly inversely with the separation of stereo features. They then argued that the planar case (where disparity varies linearly) is ineffective because it gives rise to no local differences. But Mitcheson and Westheimer (1984) pointed out 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 Mitcheson 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, and that slant percept dictated that the two vertical sides of the plane were equidistant from the viewer; later we will examine cases where the monocular configuration suggests a strongly slanted surface, with the depth relationship is again dictated by the monocular impression of slant.

Reviewing experiments 1 and 2, therefore, the common thread of explanation is that when the stimulus is simple, as in the presentation of merely four dots or two intersecting lines, the visual system interprets their depth relationships on the basis of the local differences in disparity. The cross was seen in depth with reasonable accuracy (as reflected by the subjects P and N judgments in experiment 1) because it provided a simple spatial arrangment with significant local differences (between nearby endpoints). But the P and N judgements in the cross were still influenced by the competing monocular interpretation of the intersection as a right angle in 3D. The grid data, in contrast, showed an overwhelming monocular influence for two reasons: the stereo information no longer presented local differences in disparity and secondly the grid evoked a much stronger 3D impression. In the slanted ellipse and rectangles of experiment 2, again the addition of more stereo information served to reduce the interpretation of depth from stereopsis because it presented uniform disparity differences, and simultaneously, the monocular configuration induced a more compelling 3D interpretation.

Mitcheson and Westheimer (1984) provide a framework for describing the detection of local surface discontinuities and curvature from stereopsis, but do not directly address the issue of interactions between mono and stereo processing. Putting the pieces together, we propose that surfaces can be seen in depth independently on the basis of stereopsis or monocular interpretations alone, but that interactions arise only when local disparity changes suggest either discontinuities in depth or local surface curvature. *Even though disparity values have the potential for dictating the ordering of surface points in depth, the visual system has not adopted the strategy of using disparity information to determine local depth ordering relationships, or slant, except across* *discontinuities.*

We next report additional evidence that supports this hypothesis. Three experiments were performed to examine interactions between stereopsis and the monocular interpretation, where subjects made judgments of relative depth ordering (between a probe point and a reference point, both perceived as lying on a surface) and local surface orientation. In these tasks we intentionally provided unlimited observation time and free eye movements, so that the depth and orientation of the perceived surfaces might include all potential contributions from stereopsis, including the slow effects reported by Gillam *et al.* (1984), monocular depth interpretations, and scrutiny with focal attention.

Experiment 2 also made quite apparent an anisotropy in the processing of stereo disparity, with much greater ability to detect depth in vertically-oriented, compared to horizontally-oriented, disparity gradients. This anisotropy has been described by Wallach and 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. This anisotropy appears in the following experiment as well.

2.3 Experiment 3: Detecting Planar versus Nonplanar Stereo Disparity Distributions

Thus far the stereograms contained stereo disparities induced by stereo projection of slanted plane, and the 3D interpretations have been found to be determined largely monocularly, even when that interpretation suggests depth relationships that are dramatically contradictory with the stereo disparities. In the following experiment we used both structured (line grids) and unstructured (random dot) stimuli, and examined both planar and nonplanar stereo disparity distributions (whose second directional derivatives are non-vanishing). The experiment also varied the orientation of the gradient, in light of the finding in experiment 2 that stimuli with vertical gradients assumed greater slant in depth and exhibited better shape constancy than those with horizontal gradients.

Method

Apparatus: The optical arrangement was unchanged from earlier experiments, but we now decoupled the computation of stereo disparities from the projection of the two half-images. A surface could be rendered either by lines across the surface or by random points, with the 3D surface coordinates projected to the 2D screen coordinates of each half-image in either orthographic or perspective projection. A vergence angle of 0° was specified for the stereo projection computation, which would have produced identical half-images, but prior to projection the screen coordinates were shifted laterally to introduce stereo disparity according to a specified function. The resulting field of stereo disparities, therefore, was independent of the surface being rendered. Various disparity functions were used, each a function of either the x-coordinate or y-coordinate of screen position. A constant disparity gradient of horizontal orientation, for instance, was introduced by varying the stereo disparity as a function of the y-coordinate; the disparity was zero along the horizontal meridian, positive above (with magnitude linearly increasing with distance above the horizontal midpoint of the screen) and negative below. This corresponded to a plane slanted with depth increasing vertically, the bottom of the image nearer than the top. Beside constant gradients (planar disparities), the stereo disparities could also be nonplanar, varying one-dimensionally according to a Gaussian ridge or edge profile (again, oriented both horizontally and vertically).

Stimuli: An unslanted plane was rendered by a square grid (lines separated by roughly 1.9°) or by random dots. The stereo pair was generated with stereo disparities computed by one of six functions: a slanted plane, Gaussian ridge or Gaussian edge oriented either horizontally (h) or vertically (v). The slanted plane (v), for example, introduced disparities that varied from 0′ at the center to ±51.2′ at left and right extremes of the field of view (at 6.4° eccentricity). It is termed (v) because it corresponds to a plane pivoted about the vertical meridian). The Gaussian ridge function introduced stereo disparities from 0′ in the periphery to -37.8′ along the ridge

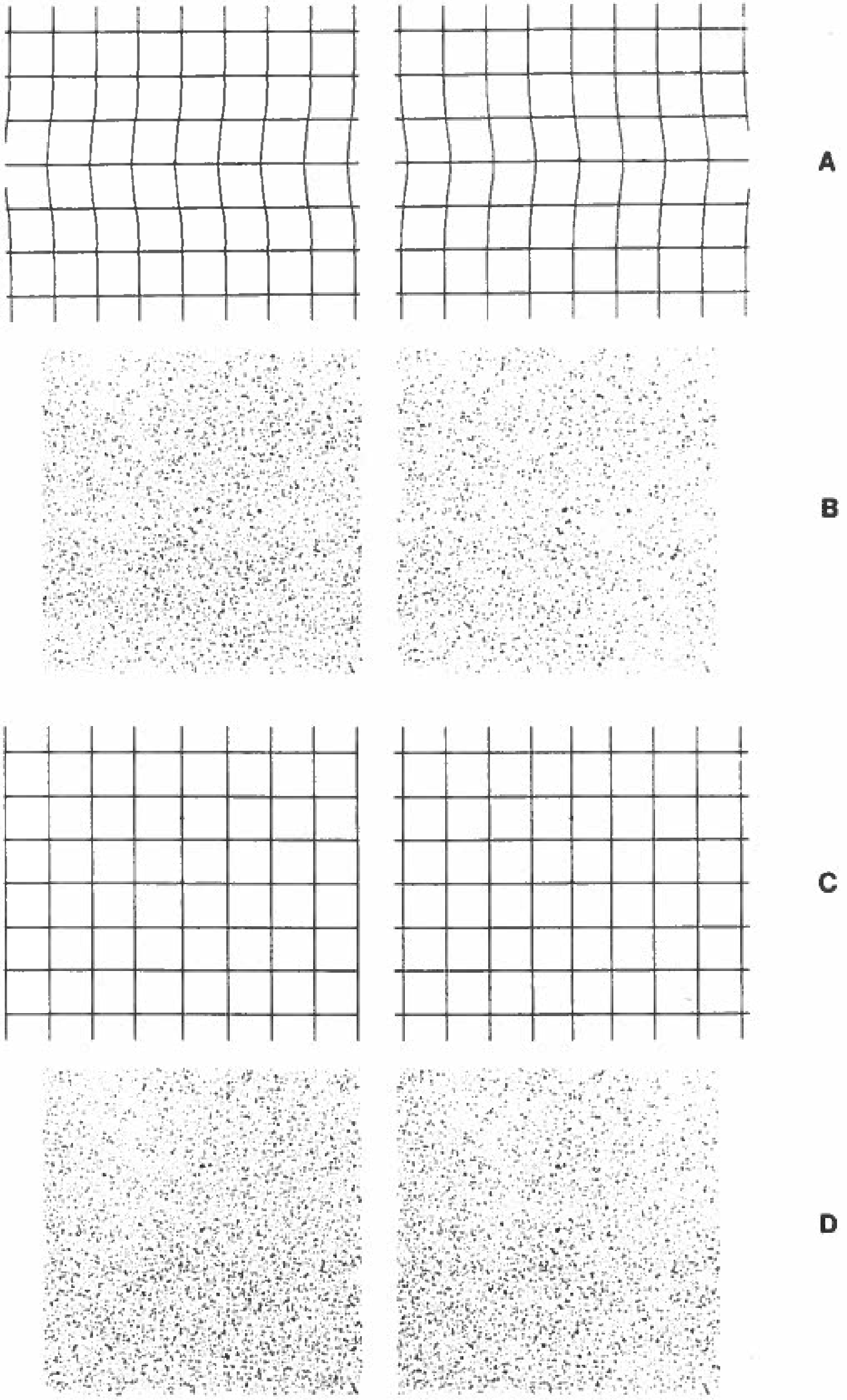

Figure 12. Horizontal Gaussian ridges in *a* and *b*; vertical Gaussian edges in *c* and d.

which was centered in the field of view. The (h) and (v) ridges ran along the horizontal and vertical meridians, respectively. The ridge was one-dimensional, protruding towards the viewer, and rather narrow (half-amplitude at ±1.6° eccentricity). The Gaussian edge, essentially one-half of the Gaussian ridge with the same space constant, varied between ±18.9′ at opposite edges of the field and passed through zero at the center. The edge presented to the viewer a moderately steep, smooth drop in depth centered in the field of view, oriented either horizontally or vertically. (Note that the constant gradient in the slanted plane presented the viewer with over twice this disparity range.) The stereo disparity functions were all continuous, since it is well known that the visual system is sensitive to even minute disparity discontinuities. The grids or dot patterns that carried these disparities extended beyond the field of view offered, to minimize edge or boundary effects. Example grid and dot stimuli are given in figures 12a and 12b (horizontal Gaussian ridges) and figures 12c and 12d (vertical Gaussian edges).

Procedure: Three subjects were used; all had excellent stereo vision and had been involved in earlier experiments. The grid and dot versions of the experiment were run separately, each with 120 trials, 5 repetitions of each of the stimuli in random order. When presented, the simulus had two large dots superimposed, both lying on the stereo surface, a reference dot at the center of the field of view and a probe dot placed at 2.9° eccentricity in one of the four cardinal directions, chosen randomly. (In the grid stimuli the probe dot always superimposed on a grid line, so as to be associated with the surface depicted by the grid.) The subject was asked to respond, by mouse interaction, whether the probe dot was closer, farther or at the same depth as the reference dot. Free eye movements and unlimited observation time were allowed. The probe and reference dots subtended 10′; the lines and dots of the stimulus were 2′ wide.

Results

Table 2 shows the responses for the grid stimuli. The data for a given stereo surface type is presented as a function of the direction of the probe dot, and indicated by a triple of numbers. Each triple gives the number of near, equidistant, and far responses for that condition. The italicized number indicates the response predicted by the stereo disparity information; the non-italicized numbers are in this sense "errors". In the case of the slanted plane (h), for example, if the probe dot was above (north of) the reference dot, it should have been seen as farther, hence the third number of the triple under N is italicized. Likewise, if the probe was east or west of the reference it should have been seen as equidistant (second number in triple). Subjects apparently ignored the fact that the radial distance to the probe increases gradually with eccentricity, and for example, regarded the probe and reference dots as equidistant when the probe was on the ridge. The nearer/equidistant/farther judgments reflected surface relief more than radial distance from the observer to the surface.

Table 2. Depth judgments for grid stimuli.

	N			S			E			W		
Slanted plane (h)	5	10	*0*	*0*	11	4	0	*15*	0	0	*15*	0
Slanted plane (v)	4	*11*	0	0	*14*	1	0	12	*3*	*0*	15	0
Gaussian edge (h)	0	0	*15*	*14*	0	1	0	*15*	0	0	*15*	0
Gaussian edge (v)	3	*12*	0	0	*13*	2	0	9	*6*	*6*	8	1
Gaussian ridge (h)	0	0	*15*	0	0	*15*	0	*15*	0	0	*15*	0
Gaussian ridge (v)	3	*12*	0	0	*14*	1	0	9	*6*	0	13	*2*

For the slanted planes (h) and (v) the data reveals that little depth was perceived from planar disparities. Specifically, none (0/30) of the N or S responses were correct for horizontal case, and only 3/30 of the E and W responses were correct for the vertical. In fact, several of the responses are the opposite of what would be expected by the stereo disparities. The performance was much better for nonplanar disparities, and generally better for the (h) than the (v) surfaces. Note that the Gaussian edge (h) and Gaussian ridge (v) data are virtually in accordance with the stereo information. The Gaussian edge (v), however, was detected with only slightly better success than the slanted plane (v).

Table 3 shows the dot pattern results. The performance was generally better, but the depth in the slanted plane stimuli was more difficult to detect than in the ridge and edge stimuli. The slanted plane (v) was the most difficult to interpret in depth.

Table 3. Depth judgments for dot stimuli.

	N			S			E			W		
Slanted plane (h)	4	0	*11*	*11*	1	3	0	*11*	4	0	*10*	5
Slanted plane (v)	6	*9*	0	0	*9*	6	0	1	*14*	*8*	1	6
Gaussian edge (h)	1	2	*12*	*15*	0	0	0	*15*	0	0	*14*	1
Gaussian edge (v)	0	*15*	0	0	*15*	0	0	0	*15*	*15*	0	0
Gaussian ridge (h)	0	0	*15*	0	0	*15*	0	*15*	0	0	*15*	0
Gaussian ridge (v)	0	*15*	0	1	*14*	0	0	0	*15*	0	0	*15*

Discussion

This experiment provides further evidence of the relative insensitivity of the visual system to planar compared to nonplanar stereo disparities, and perhaps for an anisotropy in the detection of stereo features. The subjects had a rather simple task, comparing the depth of two stereo dots, which could be performed, seemingly, by directly comparing the disparities registered at the two dots, independent of the background grid or dot pattern. The disparity difference for the slanted plane at the two locations was 22.9′, and the separation was only 2.9°. In a pilot experiment where only the probe and reference dots were presented, the task of comparing their depths was trivial and generated perfect "error free" responses. But in this experiment the probe and reference dots, while perhaps at different disparities, were at the same disparity as the grid or dot pattern in the immediate vicinity. That is, the probe and reference dots were features on the stereo surface, not free-floating, and they assumed the depth of the surface. And, if the surface was seen as flat in the image plane, then the dots, which were seen as on the surface, appeared equidistant, regardless of superthreshold differences in stereo disparity between the two dots. (It was for this reason we chose to place the dots on the grid lines, for otherwise subjects could ignore the grid and simply compare the "depths" of the dots.)

The random dot stimuli were designed to provide a baseline performance level for detecting stereo surface features in the absence of monocular structure. Given the large ranges of stereo disparity in the stimuli, we expected essentially perfect performance, and for the nonplanar disparities, that was the outcome. But the constant disparity gradients, the slanted plane stimuli, exhibited more errors than we expected, particularly for the vertical case. There was seemingly a "regression to the frontal plane" that reduced the depth differences between the probe and the reference dots. The random dot pattern clearly defined a surface, that surface had little apparent slant, and that seemingly caused the errors. Comparing the grid results with the random dot results, there appears to be an increased tendency to judge the surface as unslanted and the overall accuracy was much lower as well (there were 62 versus 37 errors). While there is a basic difficulty in perceiving depth from planar stereo disparities, as suggested by the random dot data, the performance is further degraded when a competing monocular (zero slant) interpretation is introduced by the square 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.

A final observation regarding this experiment is that any significant nonlinearity in stereo disparities seems to introduce an appropriate topographic depth feature such as a ridge or edge, into the perceived surface. The square grid appeared warped or draped over this feature. Of course, the square grid is not particularly inconsistent with the depth feature suggested by the disparity nonlinearity.

2.4 Experiment 4: Judgments of Surface Orientation with Contradictory Monocular and Stereo Information

Experiments 1 and 2 suggested that constant stereo disparity gradients had little influence on the perception of 3D over that suggested monocularly. Shape and spatial orientation in these stimuli seemed largely determined by the monocular configuration, and experiment 3 provided further evidence against our sensitivity to the planar disparities. In experiment 4 we provided stimuli that, monocularly, were strongly suggestive of a planar grid slanted in depth. The spatial orientation of the plane was quite well determined monocularly. Our question was whether the introduction of a competing constant disparity gradient influences the apparent orientation of the plane. To test that, we presented a grid and had subjects adjust a needle in 3D so that it appeared to align with the surface normal, *i.e.* to project perpendicular to the plane of the grid.

Method

Stimuli: The stimulus surface was a 4x4 grid with the same dimensions and appearance as in experiment 1. The grid was oriented randomly at two slants, 35 and 70°, and two tilts, 40 and

140° (measured counterclockwise from the positive x direction). By convention, four cardinal directions N, S, E, and W were laid out on the surface such that the north direction corresponds with the monocular tilt direction, *i.e.* the distance to the surface increases from south (bottom) to north (top), and surface points to the left (west) and right (east) lie at roughly the same distance. The surface thus increased in depth, monocularly, in either the 40° or 140° direction in the image. The stereo information, however, was consistent with distances increasing not from north to south, but increasing either to the east or to the west. That is, the monocular depth gradient and the stereo depth gradient were precisely orthogonal. To achieve this, for each of these four stimulus conditions the half-images were projected in perspective with 0° vergence angle, then the stereo disparities were computed by the slanted plane (constant gradient) disparity function used in experiment 3, with the direction of the disparity gradient oriented to be ±90° away from the specified tilt. While monocularly the normal should point to the north, the stereo information implied the surface normal should point to the east or west (depending on the polarity of the disparity gradient). The 8 conditions (four monocular surface orientations times two stereo disparity gradients) were run on the 3 subjects with 5 trials per condition.

Procedure: The three subjects used in experiment 3 participated in this experiment as well. A stimulus surface was presented, and superimposed in stereo at the center of the pattern was a needle (described earlier) that could be adjusted in 3D to point in any direction. One end of the needle appeared fixed at the center intersection point of the grid, and to extend in 3D above the plane of the grid. With mouse interaction the subjects could step the needle in tilt (orientation in the image plane) and slant (orientation perpendicular to the image plane) so that it appeared perpendicular to the grid plane. Increments of 2.5° in slant and tilt were regarded to provide sufficiently fine grain for this task. Adjusting the normal in 3D was an immediate and natural task for the subjects. For each trial the subject adjusted the normal until satisfied that it appeared three-dimensionally perpendicular to the grid, then advanced to the next stimulus. Unlimited time was allowed for each judgment.

In a second set of presentations, the experiment was repeated with the same stimulus set, but the needle that was adjusted by the subject to appear perpendicular to the surface was monocularly presented, to the dominant eye only (figure 13). The line could be rotated about the center point of the grid, changing its tilt until it appeared to point in the direction of the normal to the surface. The three subjects, having completed the experiment with the stereoscopic probe, noted that the monocular probe appeared much less three-dimensional, as would be expected, but were able to make the tilt judgments with confidence nonetheless. The subjects had no difficulty observing the surface binocularly with the probe superimposed on only one half-image. The probe appeared fixed to the surface and appeared to pivot in 3D about the fixed end. When adjusted to appear normal to the grid it attained a surprising degree of depth. The three subjects were presented 5 trials per each of the eight conditions, as before.

Results

Table 4 presents the mean surface orientations resulting from this task. This data can be examined for influence of the stereo disparity gradient on either the apparent tilt or slant of the monocular surface. Since the disparity gradient was orthogonal to the monocular depth gradient, the normal could be expected to lean in the direction of the disparity gradient. That is, the tilt should rotate counterclockwise when the disparity gradient was to the west, and clockwise when the gradient was reversed to the east. There was a clear trend in the tilt data; for each combination of slant and tilt the mean tilt was significantly greater when the gradient was to the west than to the east (at the 0.02 level or lower). The apparent tilt was systematically rotated in the direction of the stereo disparity gradient.

Figure 13. The disparity gradient is perpendicular to the apparent monocular gradient of depth. Subjects adjusted the monocular "normal" until it appeared perpendicular to the grid in 3D.

Table 4. Surface orientation judgments with stereoscopic normal.

Surface Orientation		Disparity Gradient to West				Disparity Gradient to East			
slant	tilt	slant	std	tilt	std	slant	std	tilt	std
35.0	40.0	36.5	(2.8)	56.2	(4.5)	37.5	(4.2)	41.0	(10.3)
35.0	140.0	33.0	(3.6)	147.0	(12.2)	38.8	(6.1)	130.3	(8.1)
70.0	40.0	68.5	(4.6)	46.8	(6.2)	64.7	(8.7)	42.0	(7.0)
70.0	140.0	65.0	(7.7)	143.2	(4.6)	66.5	(6.6)	135.7	(4.8)

In contrast, apparent slant was not significantly influenced by the stereo disparity gradient. The means across reversals in the direction of the stereo disparity gradient did not differ significantly except for one condition (slant = 35°, tilt = 140°, t(28) = 3.19, $p < .01$). The mean apparent slant was also within one standard deviation of the monocular slant (either 35° or 70°) for all conditions. In fact, for (slant = 70°, tilt = 40°) the mean apparent slant of 68.5° was not significantly different from 70° at even the .2 level (t(28) = 1.26).

Table 5. Surface tilt judgments with monocular normal.

Surface Orientation		Disparity Gradient to West		Disparity Gradient to East	
slant	tilt	tilt	std	tilt	std
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)

In the second series of presentations with the monocular normal, the tilt judgements were much less influenced by the stereo disparity gradient. Of the four surface orientation conditions in only one case was the mean apparent tilt significantly different for the two directions of the disparity gradient (slant = 70°, tilt = 40°, t(28) = 2.41, p = .02). The apparent tilt was generally close to the monocular tilt.

Discussion

The apparent slant of the grid surface was not influenced by the constant disparity gradient, and corresponded closely with that indicated by the monocular perspective image. The subjects uniformly saw the grid as oriented according to the monocular depth gradient. The tilt component showed some "leaning" in the direction of the disparity gradient, however. That is, subjects tended to rotate the normal toward the direction of increasing depth according to the stereo disparities. We suspected that this was due to the fact that the normal was presented in stereo, and there was an interaction between the normal and the grid lines in the immediate vicinity. For that reason we chose to also probe the apparent tilt monocularly, by a technique used earlier on similar stimuli (Stevens, 1983). When the subjects were to visualize whether the tilt of the monocular line corresponded with the direction the surface normal would project, the results showed a close correspondence between apparent and monocular tilt. We conclude that the perspective (effectively monocular) grid image was backprojected into 3D with little concern for the fact that there was a constant gradient of stereo disparity that suggested a contradictory spatial orientation.

The stereo disparity gradient had some effect, however. The pattern often appeared rather flat in the image plane, and although subjects readily indicated the slant and tilt of the surface normal (and hence the direction of increasing distance on the surface) there was generally little variation in depth — as Gillam (1968) observed in reversed-disparity stereograms of perspective projections. The apparent slant was close to that specified by the monocular projection, but as is offen the case in viewing a monocular image, the apparent depth was not that compelling. That is certainly not surprising, since the stereo disparity was constant in that direction. The stereo disparity was oriented perpendicular to the monocular depth gradient, recall, and that should have given a clear suggestion of distance increasing in that direction. But as found in experiments 1-3, this gradient of disparity was ineffectual. At this point we applied the experimental approach in experiment 3, comparing the apparent depth at a probe and reference points, to further confirm this rather unexpected finding.

2.5 Experiment 5: Judgments of Depth with Contradictory Monocular and Stereo Information

In this final experiment we tested the effect of contradictory monocular and stereo depth gradients on the ability to judge depth at points on the surface.

Method

Stimuli: The grid stimuli were the same as used in experiment 4, with slant fixed at 65° and two tilts, 45 and 135°. As before, the four cardinal directions were laid out on the surface such that north corresponds with the monocular tilt direction. The stereo disparities were computed according to a linear disparity function; the direction of the disparity gradient was aligned at random with one of the cardinal directions. When the disparity gradient was northward they were consistent with the monocular gradient of distance. When the gradient increased to the east or west it was orthogonal to the monocular perspective (as in the previous experiment), and when south the stereogram had effectively reversed disparities.

Procedure: Four subjects were used. All had previous experience in the experimental series. A stimulus surface was presented, with a dot superimposed in stereo 6° from the center in either the north, south, east, or west direction. The subject was instructed to respond, via the mouse, whether the surface, at the position of the dot, was nearer, farther or at the same depth as a reference point on the surface at the center of the viewing screen (figure 14). There were 32 stimuli, 2 tilts (45° and 135°), 4 probe locations (N, S, E, W), and 4 directions for the disparity

---

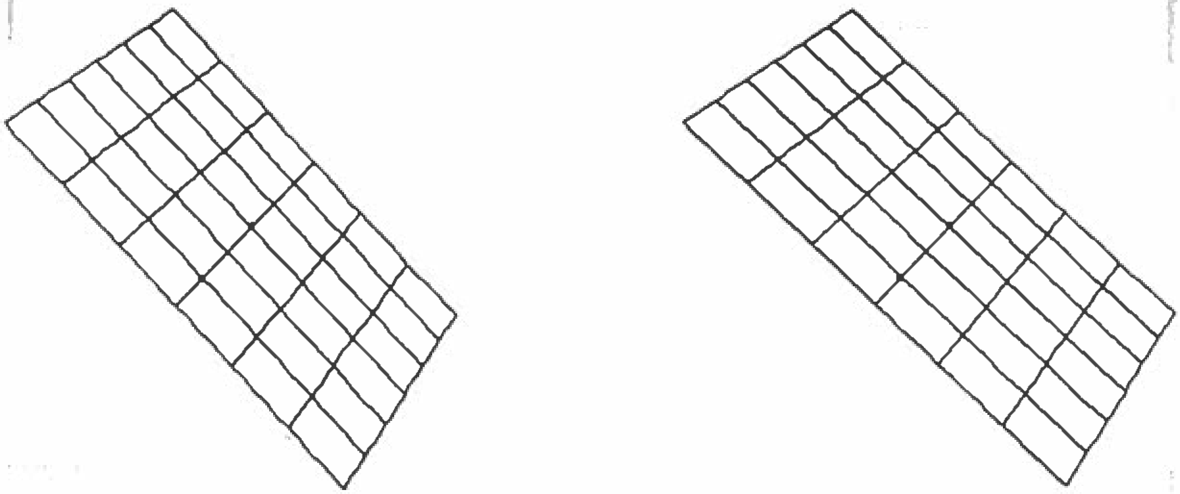

Figure 14. As in figure 13, the disparity gradient is perpendicular to the apparent monocular gradient of depth. In this experiment subjects compared the relative depth of the probe point with the central reference point.

gradient. The experiment consisted of 5 repetitions of the 32 stimuli in random order.

Results and Discussion

Table 6 shows the sets of closer/equidistant/farther responses for each combination of probe location (across) and disparity gradient direction (down). Again, the italicized values indicate the responses predicted by the stereo disparity. The first row is a control, since the direction of the stereo and mono gradients coincide (mono depth and disparity increase towards the north). Note that when the probe was to the east or west there was a greater tendency to see the probe and reference points as equidistant (21/40 and 22/40 trials). As expected, in 40/40 trials the probe was seen as farther when north of the reference and as nearer when south of the reference.

Table 6. Depth judgments for perspective grid (monocular gradient to North).

Disparity gradient	N			S			E			W		
North	0	0	*40*	*40*	0	0	10	*21*	9	7	*22*	11
South	*1*	5	34	37	3	*0*	3	*27*	10	9	*28*	3
East	0	*0*	40	40	*0*	0	13	17	*10*	*7*	24	9
West	0	*5*	35	35	*5*	0	*7*	24	9	9	25	*6*

When disparity was reversed (see second row) there was an overwhelming tendency to continue to see the probe as farther (34/40 trials) when the probe was to the north, and (37/40) when to the south. Slight "regression to the frontal plane" was evident in the fact that 7/80 trials (combination of probe to the north and to the south) resulted in "equidistant" judgments. In only one trial was the depth actually judged according to the stereo information. Subjects seemed to have difficulty deciding the relative depth of the probe and reference points in the E and W probe locations, hen they were approximately equidistant, even in the case where stereo and perspective were consistent (first row of table 6). The E and W data for all directions of the disparity gradient were quite similar, showing no systematic bias according to the direction of the disparity gradient. Even when the stereo disparity gradient was orthogonal to the monocular depth gradient (lower two rows of table 6) the relative depth judgments were dominated by the monocular interpretation. For probe locations N and S there was virtually no regression, despite the fact that the probe and reference points were at zero disparity. Likewise, for probe locations E and W there was no apparent bias as a function of stereo disparity direction, with most trials judged "equidistant". As consistent with the tilt results in experiment 4, the orthogonal disparity gradient had negligible influence on the perceived depth orderings.

Overall, the apparent depth corresponded very closely with that given by the monocular information. Even in the case where the stereo gradient was in the opposite direction to the monocular depth gradient the monocular interpretation prevailed. The disparity gradient had little effect against the contradictory monocular perspective information. Gillam (1968) found that when perspective and stereo are in opposition compromise slant judgements result. We find that when the two are orthogonal there is likewise a slight diminution of apparent slant, but not enough to influence the apparent depth ordering.

3. General Discussion

This series of experiments has shown that the monocular 3D interpretation of planar surfaces is influenced very little by linear stereo disparity distributions. In part, the findings underscore recent observations by Gillam *et al.* (1984) and Mitcheson and Westheimer (1984) regarding the weak salience of constant disparity gradients. It also reveals the somewhat surprising fact that even though stereopsis could in principle establish depth ordering relationships across a slanted plane, it does not. We conclude from this that the computational goals of stereopsis, vis-a-vis surface perception, are not the determination of depth but the detection of surface discontinuities and curvature. Stereopsis of course also provides distance information, but that seems to serve a very different goal, namely the localization of objects in space, not the local description of surface shape and disposition. We found little evidence that slant (or equivalently, depth gradient direction or magnitude) is computed locally from stereo disparity in our stimuli.

In summary, we sugggest that there are possibly three distinct methods for analyzing stereo disparity information. The first concerns the extraction of surface curvature assertions (topographic features), which corresponds with Gillam *et al.*'s (1984) "boundary mode" processing. We extend their notion by pointing out that it is not merely sharp discontinuities in disparity that drive stereo depth, but disparity distributions with non-vanishing second spatial derivatives. More than just boundary mode processing, we regard it as surface curvature processing, with boundaries constituting a limiting case of extreme curvature. That is, in addition to the sharp disparity discontinuities that Gillam *et al.* (1984) examined (analogous to sharp intensity edges), the visual system interprets a variety of systematic variations in stereo disparity in terms of the corresponding surface curvature features, by methods analogous to the detection of edge, bar, blob and other intensity features in a retinal image. The stereo system is essentially a "depth edge detector", by our view. While it also has been calibrated in conjunction with convergence and motion parallax, to provide range information at specific locations, under scrutiny and focal attention, it is remarkable that this range system is limited in its ability to interact with surface perception, nonetheless.

The second method of stereo processing Gillam *et al.* (1984) term "surface mode" is a slow process that gradually develops depth. While that mode of depth filling occurs, it seems to be strongly constrained by the interpretation derived by the faster "boudary mode" system and

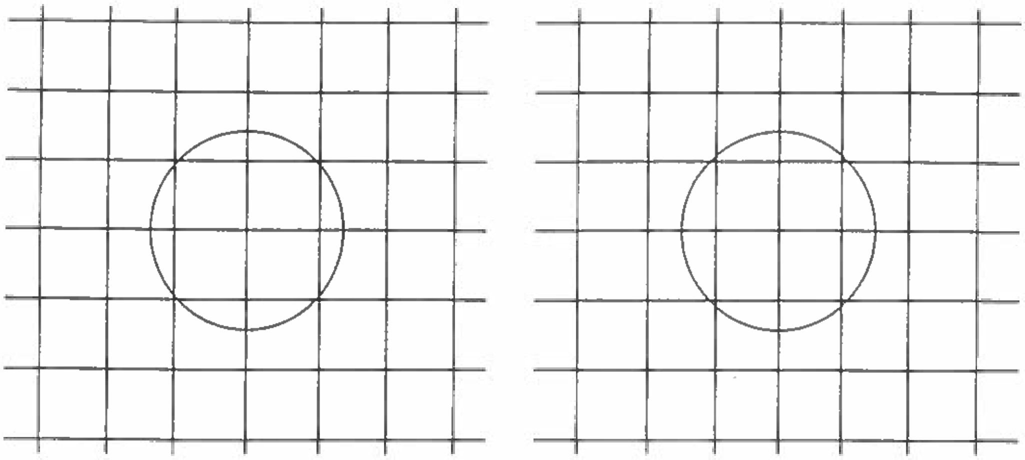

Figure 15. Stereoscopic analogue to the brightness contrast effect.

significantly, by the monocular interpretation.

The third mode of stereo processing is the seemingly independent determination of absolute distance, or range, information from accomodation, vergence, motion parallax and stereo disparity (Johansson 1973). Ritter's (1979) observation that perceived depth is most precise for small disparity values around a given vergence angle, implies that the convergence angle is one of the main determiners of range. We suspect that the perception of absolute distance primarily subserves motor functions, including locomotion and manipulation. The fact that it is most accurate only up to approximately 2 m is also suggestive.

We have also extended the analogy to intensity edge detection, first characterized as a depth version of the Craik-O'Brien-Cornsweet illusion (Anstis, Howard & Rogers, 1978; Rogers & Graham, 1983; Gillam, Flagg & Finlay, 1984). We have shown that stereo disparity contrast induces depth contrast in a manner closely analogous to brightness contrast effects (Ratliff, 1965) observed in the luminance domain (Stevens & Brookes, 1986c). A circle of constant 0° disparity viewed against a background of constant disparity gradient is seen as slanted in depth in the direction opposite to the background gradient. The effect is quite robust, and was demonstrated for a variety of stimuli in which the depth was related to the local disparity contrast, not the absolute disparity, in quite the same way as brightness is related to the local luminance contrast and not the absolute luminance. Figure 15 shows a stereogram of a circle of zero slant embedded at zero disparity within a stereogram of a Gaussian edge that passes through zero disparity at the vertical meridian. Note that the circle assumes a complementry slant in depth, induced by the local disparity contrast. Likewise, figure 16 shows the slanted rectangles stimulus from experiment 2, which tends to appear in depth as a corridor, with a circle of constant zero disparity superimposed. The circle assumes a complementary slant and the local disparity contrast enhances the perception of the rectangles as slanted. We seek to extend this observation, taking into consideration the proximity effects observed by Mitcheson and Westheimer (1984), in order to model the notion of "local" with regard to disparity contrast. We are intrigued with the prospect

Figure 16. The zero-disparity circle assumes a complementary slant to the slanted rectangles and tends to diminish the corridor-in-depth interpretation.

that the visual system processes stereo disparities much as it does retinal intensities. Just as we are highly sensitive to non-linear changes in intensity (such as edges and bars over a range of scales or spatial frequencies) and yet relatively insensitive to slow, linear or ramp-like intensity changes, we are acutely sensitive to non-linear, but not linear, changes in stereo disparity. In terms of retinal neurophysiology, it is well recognized that linear or constant intensities do not stimulate retinal ganglion cells because of the balanced lateral-inhibitory nature of their receptive fields. Computationally, it would be said that these mechanisms are computing Laplacian-like convolutions of the intensity signal, for which only those intensity distributions having a non-zero second directional derivative are "detected" (Marr & Hildreth, 1980). It would be exciting to find analogous differentiation operators whose receptive fields summate raw disparity values.

Finally, with regard to the integration of 3D information proper, we conclude that monocular 3D interpretations are much more important than previously expected. Despite their intrinsic ambiguity of interpretation, the visual system relies on monocular configurations for hypothesizing 3D arrangements and spatial dispositions. Figure 17 shows concentric ellipse that are slanted 60° to the line of sight, as in experiment 2. But now the ellipses have 3:1 aspect ratio, which leads to three distinct interpretations in 3D. First, the stereo information suggests quite literally a slanted plane. The second interpretation, as before, is of a tunnel in depth, but this time of elliptical cross section. The third interpretation is of a downward view of a flat circular platter, where the elliptical form is attributed to slant. The observer can, at will, switch among these three radically different 3D interpretations.

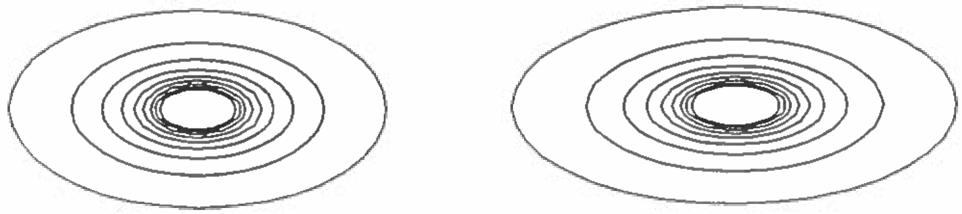

Figure 17. The slanted plane of concentric ellipses in stereo can also be interpreted in 3D monocularly as either a view down an elliptical tunnel or as a downward view of a flat circular platter.

4. REFERENCES

Anstis, S., Howard, I.P. & Rogers, B. 1978 A Craik-O'Brien-Cornsweet illusion for visual depth. *Vision Research* **18**, 213-217.

Attneave, F. & Frost, R. 1969 The determination of perceived tridimensional orientation by minimum criteria. *Perception and Psychophysics* **6**, 391-396.

Dosher, B.A., Sperling, G. & Wurst, S.A. 1986 Tradeoffs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure. *Vision Research* **26** 973-990.

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

Gillam, B.J., Flagg, T. & Finlay, D. 1984 Evidence for disparity change as the primary stimulus for stereoscopic processing. *Perception and Psychophysics* **36**, 559-564.

Gogel, W.C. 1956 Relative visual direction as a factor in relative distance perceptions. *Psychological Monographs* **418**, 1-19.

Gogel, W.C. 1972 Depth adjacency and cue effectiveness. *Journal of Experimental Psychology* **92**, 176-181.

Gogel, W.C. & Mershon, D.H. 1977 Local autonomy in visual space. *Scandinavian Journal of Psychology* 18, 237-250.

Gregory, R. 1970 *The intelligent eye.* London: Weidenfeld and Nicolson.

Gulick, W.L. & Lawson, R.B. 1976 *Human stereopsis: A psychological analysis.* New York: Oxford University Press.

Johansson, G. 1973 Monocular movement parllax and near-space perception. *Perception* **2**, 135-146.

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

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. & Hildreth, E. 1980 Theory of edge detection. *Proc. Roy. Soc. Lond. B.* **207**, 187-217.

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 Research* **23**, 191-198.

Mitchison, G.J. & Westheimer, G. 1984 The perception of depth in simple figures. *Vision Research* **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. & Mizraji, E. 1985 Errors in the stereoscopic separation of surfaces represented with regular textures. *Perception* **14**, 315-328.

Ono, H. & Comerford, J. 1977 Stereoscopic depth constancy. In *Stability and constancy in visual perception: Mechanisms and processes*, W. Epstein, ed. New York: Wiley.

Perkins, D.N. 1972 Visual discrimination between rectangular and nonrectangular parallelopipeds. *Perception and Psychophysics* **12**, 396-400.

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

Ratliff, F. 1965 *Mach bands*. San Francisco: Holden Day.

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

Ritter, M. 1977 Effect of disparity and viewing distance on perceived depth. *Perception and Psychophysics* **22**, 400-407.

Ritter, M. 1979 Perception of depth: Processing of simple positional disparity as a function of viewing distance. *Perception and Psychophysics* **25**, 209-214.

Rogers, B.J. & Graham, M.E. 1983 Anisotropies in the perception of three-dimensional surfaces. *Science* **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*, M. Kubovy and J. Pomerantz, eds. Hillsdale, N.J.: Lawrence Erlbaum.

Smith, A.H. 1965 Interaction of form and exposure time in the perception of slant. *Perceptual and Motor Skills* **20**, 481-490.

Smith, O.W. & Smith, P.C. 1957 Interaction of the effects of cues involved in judgements of curvature. *American Journal of Psychology* **70**, 361-375.

Smith, P.C. & Smith, O.W. 1961 Vericial perceptions of cylindricality: a problem of depth discriminatio and object identification. *Journal of Experimental Psychology* **62**, 145-152.

Stevens, K.A. 1983 Surface tilt (the direction of slant): a neglected psychophysical variable. *Perception and Psychophysics* **33**, 241-250.

Stevens, K.A. 1983a Slant-tilt: The visual encoding of surface orientation. *Biological Cybernetics* **46**, 183-195.

Stevens, K.A. 1983c Tilt (the direction of surface slant): a neglected psychophysical variable. *Perception and Psychophysics* **33**, 241-250.

Stevens, K.A. & Brookes, A. 1986c A brightness contrast analogue in stereopsis. *Perception*, in preparation.

Ullman, S. 1979 *The interpretation of visual motion*. Cambridge, MA: MIT Press.

Wallach, H., Gillam, B., and Cardillo, L. 1979 Some consequences of stereoscopic depth constancy. *Perception and Psychophysics*, 235-240.

Wallach, H. & Bacon, J. 1976 Two forms of retinal disparity. *Perception and Psychophysics* **19**, 375-382.

Wallach, H. & Zuckerman, C. 1963 The constancy of stereoscopic depth. *The American Journal of Psychology* **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**(3), Part II, 504-523.

Yellott, J.I. & 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 Research* **16**, 79-82.

Appendix: Equations Relating Stereo Disparity and Angle of Intersection

The stereo image of two intersecting line segments $\vec{U}$ and $\vec{V}$ contains information about their angle of intersection β in 3D. One may derive an expression $\beta = f\ (d, I, \delta_u, \delta_v)$ on the basis of stereo disparities, δ_u and δ_v, and the viewing distance d and interpupillary separation I.

Consider the two unit vectors, $\vec{U}$ and $\vec{V}$, embedded in a Cartesian x-y-z system, where the z-axis corresponds to the view vector and the image plane is parallel to the x-y plane.

$$\vec{U} = \left\{ \cos \alpha_u , \sin \alpha_v , U_z \right\}$$

$$\vec{V} = \left\{ \cos \alpha_v , \sin \alpha_v , V_z \right\}$$

Since $\vec{U} \cdot \vec{V} = |\vec{U}|\,|\vec{V}| \cos \beta$, we have that $\beta = \cos^{-1} \dfrac{\vec{U} \cdot \vec{V}}{|\vec{U}|\,|\vec{V}|}$,
where

$$\vec{U} \cdot \vec{V} = (\cos \alpha_u \cos \alpha_v + \sin \alpha_u \sin \alpha_v + U_z V_z) = (\cos(\alpha_u - \alpha_v) + U_z V_z)$$

$$|\vec{U}| = (\cos^2 \alpha_u + \sin^2 \alpha_u + U_z^2)^{1/2} = (1 + U_z^2)^{1/2}$$

$$|\vec{V}| = (\cos^2 \alpha_v + \sin^2 \alpha_v + V_z^2)^{1/2} = (1 + V_z^2)^{1/2}$$

Regarding the stereo projection, assume $|\vec{U}|$, $|\vec{V}| << d$, i.e. one observes only a local region of the surface, such that the vector lengths are small compared to the viewing distance. We may relate the z (depth) components of the two vectors to their disparities U_z and V_z by

$$U_z = \frac{d^2 \delta_u}{I}$$

$$V_z = \frac{d^2 \delta_v}{I.}$$

Combining the above we solve for β:

$$\beta = \cos^{-1} \left[\frac{\cos(\alpha_u - \alpha_v) + \dfrac{d^4}{I^2} \delta_u \delta_v}{\left(1 + \dfrac{d^4}{I^2} \delta_v^2\right)^{1/2} \left(1 + \dfrac{d^4}{I^2} \delta_u^2\right)^{1/2}} \right] \tag{1}$$