Biol. Cybern. 56, 355–366 (1987)

Biological
Cybernetics

Probing Depth in Monocular Images*

Kent A. Stevens and Allen Brookes

Department of Computer Science, University of Oregon, Eugene, OR 97403, USA

Abstract. It is generally expected that depth (distance) is the internal representational primitive that corresponds to much of the perception of 3*D*. We tested this assumption in monocular surface stimuli that are devoid of distance information (due to orthographic projection and the chosen surface shape, with perspective projection used as a control) and yet are vividly three-dimensional. Slant judgments were found to be in close correspondence with the actual geometric slant of the stimuli; the spatial orientation of the surfaces was perceived accurately. The apparent depth in these stimuli was then tested by superimposing a stereo depth probe over the monocular surface. In both the perspective and orthographic projection the gradient of perceived depth, measured by matching the apparent depth of the stereo probe with that of the monocular surface at a series of locations, was substantial. The experiments demonstrate that in orthographic projection the visual system can compute from local surface orientation a depth quantity that is commensurate with the relative depth derived from stereo disparity. The depth data suggests that, at least in the near field, the zero value for relative depth lies at the same absolute depth as the stereo horopter (locus of zero stereo disparity). Relative to this zero value, the depth-from-slant computation seems to provide an estimate of distance information that is independent of the absolute distance to the surface.

1 Introduction

There have been few experimental studies in which apparent depth is measured directly. Gregory (1968, 1970) developed an apparatus (the so-called "Pandora's box") whereby a stereoscopically-viewed probe could be superimposed over a monocularly-presented stimulus to measure apparent depth (see e.g. Deregowski 1970; Kennedy 1974; Chevrier and Delorme 1983; Topper and Simpson 1981). The fact that subjects can match the depth of a stereoscopically-viewed dot with that perceived in a monocular image is actually remarkable. The monocular image is apparently localized in the same space as the binocular probe, and is perceived as extending in depth in a manner that can be compared with the depth of the probe. Usually the ability to measure distance in a monocular stimulus is taken for granted; this ability, however, reveals a basic strategy for integrating monocular and binocular depth perception.

Consider the perspective-projected surface suggested in Fig. 1a and the orthographic projection of that surface in Fig. 1b. These figures provide an immediate impression of three-dimensionality, which is significantly enhanced when they are viewed monocularly as luminous lines against a featureless background. One can readily judge local surface orientation across the depicted surface (the orientation to the tangent plane to the surface relative to the observer), its topography (the folds, troughs, and ridges, and so forth), and seemingly, perceive variations in depth across the surface. In the orthographic image, the perceived surface shape – its spatial orientation and topography – is due to the geometry of the image curves. Since the 2*D* configuration might correspond to any of an infinity of possible 3*D* surface configurations, and yet observers see a particular 3*D* shape, there are apparently strong constraints imposed on the 3*D* interpretation. One theoretical suggestion for the constraints is that the visual system presumes the image curves correspond to lines of principal curvature across the apparent surface (Stevens 1981, 1983, 1986). This geometrical constraint (plus constraint afforded by assuming representative or generic viewpoint and

---

* Supported by Office of Naval Research Contract N00014-K-84-0533. We gratefully acknowledge the suggestions of Jacob Beck regarding the experimental design, and the assistance provided by Cathryn Stanford

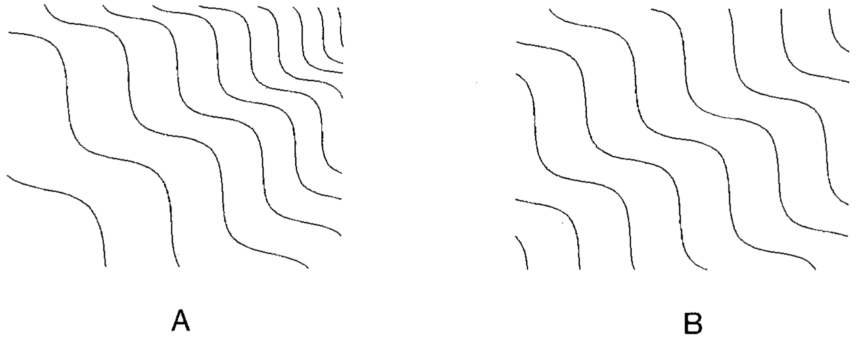


Fig. 1A and B. In **A** a sinusoidal cylinder is rendered by a family of parallel contours, specifically lines of curvature, in perspective projection. In **B** the same surface is seen in orthographic projection

contour placement) allows one to infer much about the shape of the corresponding surface topography and spatial orientation. But since the projected scale of the surface is constant across the image, there is no information about distance, either relative or absolute, available in the image. Thus, to the extent that depth is perceived in orthographic projections such as these, it must be derived indirectly from perceived local surface orientation.

The experiments discussed below suggest that depth does vary in a quantitative and continuous manner across orthographic projections that are devoid of distance information. They also raise the question of how the monocular depth is brought into register with that derived from the stereo probe. The experiments were designed so that the task required comparing the apparent depth of the stereo probe with the depth of the monocular surface. To establish the apparent *3D* orientation of the surface, an initial experiment was performed in which subjects adjusted a slant probe at various locations across the perceived surface.

2 Experiments

2.1 Slant Experiment

Method

In this experiment subjects viewed the surfaces shown in Fig. 1. The surface was presented monocularly yet appeared vividly in *3D*, and the task of this experiment was to quantify the apparent slant at locations across this surface. Apparent slant was measured by a small ellipse of variable aspect ratio (ratio of minor to major axis length) that appeared at one of four locations on the surface. It is well known that an ellipse of given aspect ratio can be interpreted in *3D* as a foreshortened circle oriented relative to the observer such that the slant (the angle between the line of sight and the surface normal) equals the cosine of the aspect ratio, and the tilt (the direction the normal in the image plane) is aligned with the orientation of the minor axis[1]. Since observers readily interpret an ellipse as a circle slanted in *3D*, it can be used as a probe of apparent slant, in which the aspect ratio is adjusted until the apparent slant of the circle matches the apparent slant of the surface on which it is visualized. The subject observes the surface prior to superimposing the ellipse, then if the ellipse is foreshortened appropriately (such that it has the same apparent slant as the underlying surface in the vicinity) it will appear to lie flush on the surface. In comparing the apparent slant of the circle and the underlying surface it is best to have the probe's other spatial degree of freedom, the tilt, geometrically correct according to the monocularly-depicted surface.

Stimuli: The surface stimuli consisted of perspective and orthographic projections of a surface rendered by lines of curvature. In all experiments we used a common surface, a singly curved, sinusoidal cylinder presented in either perspective or orthographic projection Fig. 1). Despite the fact that only a family of parallel curves was displayed, subjects readily interpreted the curves in *3D* as lying on an imagined smooth surface. We used the spaces between the contours for projecting the slant and depth probes; there was an adequate impression of a smooth surface extending in space spanning the parallel curves to support such judgments. In a given trial the stimulus surface was either presented as shown in Fig. 1 or mirror-reflected, at random, to balance between stimuli where depth increased towards the right and towards the left. The stimuli were generated by a Symbolics 3600 Lisp Machine and, in this slant experiment, were projected on a Tektronix 690SR monitor as luminous lines (in

---

[1] Note that the ellipse is ambiguous in depth, and while the normal would project in the orientation of the minor axis, the normal might point in one of two directions separated by 180°. This ambiguity is shared by the orthographic surface stimuli, of course, but the stimuli were oriented to favor the tendency to interpret depth as increasing upwards in the image. Observers saw the orthographic stimuli unambiguously in depth, and the ellipse was seen correspondingly

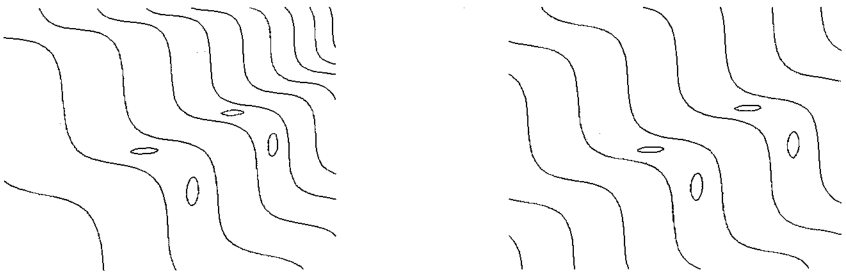


Fig. 2. The ellipses represent the geometrically correct local surface orientations for the projections four slant probe locations for the perspective and orthographic projections

the later depth experiments, the stimuli were projected on a pair of monochrome monitors). The subjects viewed the surface stimuli in a darkened room from approximately 2 m; the surfaces subtended roughly 6° of visual angle.

Procedure: Four graduate students participated as paid subjects; all were naive to the purposes of the experiments. Observers were asked to compare the apparent slant of the surface with that of the slant probe, which consisted of an ellipse that subtended approximately 30′ with the minor axis oriented to coincide with the surface tilt at the given probe location (see Stevens 1983a, 1983b). The probe was sufficiently large to have a clearly perceived slant, but small enough to appear to lie flush with the undulating surface when properly slanted. The primary experimental variable was the degree of foreshortening of the ellipse, which implied a corresponding slant of the probe, when seen as a circle in 3*D*. The subjects were to compare the apparent slant of the circle with that of the surface at the given probe location.

A randomized-staircase forced-choice method was used to converge the slant of the probe to coincide with that of the stimulus surface at four locations across the surface. The probe locations used in the experiment are shown in Fig. 2; the geometrically-computed slant for the orthographic projection was approximately 78° at the upper two locations and 65° at the lower.

The experiment was initialized with the slant probe at each location purturbed randomly by 10°. A stimulus trial consisted of, at random, either orthographic versus perspective projection, and one of the four probe locations, and either mirror-reflected or not. A trial began with the surface presented for 500 ms during which time the subject was to visualize the surface in 3*D* while holding central fixation. The slant probe was then superimposed over the stimulus surface, and the combined display was presented for an additional 750 ms during which the subject was to decide whether the slant of the probe was greater or less than that of the surface at that location. These judgements were recorded by pushing the left or right button of a mouse. Subjects were allowed to respond with the middle button if uncertain, whereupon that trial was not recorded. After the additional 750 ms had expired, or after responding, the display was blanked. The subject's response initiated the presentation of the next trial. Each slant judgment changed the slant of the probe by 5° appropriately for the next presentation of the probe at that probe location. The experiment continued until the subject had sent the probe "through" the apparent slant 10 times at each of the four probe locations.

Results and Discussion

The mean slant judgements across the four subjects were analyzed for each probe location, for both perspective and orthographic projection. In the perspective stimuli, the differences between the mean apparent slant and the geometrically-computed slant were small, varying from an underestimation by 4.8° to an overestimation by 0.6°, with standard deviations of no more than 3.9°. The mean difference between apparent and geometric slant, averaged across subjects and probe locations, was only −0.69° (underestimation).

Similarly, for the orthographic stimuli, the differences between apparent and geometric slant varied from underestimation by 4.15° to overestimation by 4.1°, with standard deviations of as much as 8.1° (twice that in the perspective case). The mean difference between the apparent and geometric slant was only −1.0°, averaged across subjects and probe locations for the orthographic stimuli. Reviewing the data for both the perspective and orthographic stimuli, the mean slant judgement at each probe location was significantly different from the associated "correct" slant in only 2 out of the 8 conditions (4 probe locations and perspective versus orthographic).

The purpose of this experiment was to quantify the apparent slant in the depth stimuli, and in doing so we found that apparent slant was in very close agreement with the geometrically correct slant, both for the perspective and orthographic case. With this observation we now asked whether the apparent depth in these stimuli reflects the very significant amount of apparent slant. While the orthographic slant judgments were more variable, there was very little systematic slant

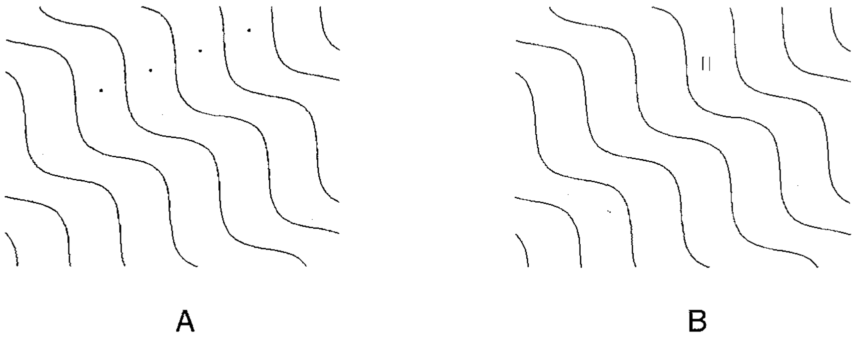


Fig. 3A and B. The four probe locations are shown in **A** (numbered 0–3 from left to right). In **B** the depth probe is depicted as a pair of bars with horizontal disparity, superimposed on the monocular surface at probe location 2

underestimation. The apparent slant was, to the precision we needed for this benchmark, essentially veridical in both projections. We were therefore interested to see how depth was perceived in these stimuli. As to whether depth would vary in the orthographic projection consistently with the apparent slant was, in our opinion, and open question. We in fact suspected that the "apparent three-dimensionality" of the orthographic case might turn out to be illusory, that depth would not vary as the slant would suggest, and that the *3D* impression reflects perceived surface orientation rather than distance.

2.2 Depth Experiments

Method

Apparatus: A Wheatstone-style stereoscope was constructed to superimpose a binocular depth probe on a monocular surface stimulus. A pair of optically-flat front-surfaced mirrors were used with two flat-screen, low geometric distortion Tektronix 634 monochrome displays. The monitor screens were positioned 79 cm from the observer (measured along the optic axis from eye to screen). In this arrangement the stimulus surface subtended approximately 7°, projected monocularly to the dominant eye. The depth at a point in this display could be measured by superimposing on this image a stereoscopic probe, much as in Pandora's box. Note that in this apparatus the binocular and monocular images were projected on the same physical CRT surfaces, and thus no "motion parallax" depth cue could be induced by head movement[2]. This obviated the need for a device to restrict the free motion of the head. The stimuli were projected as luminous lines against a dark background. The stereoscope was shielded from ambient light, and the left and right monochrome monitors were driven independently by the red and green channels of a color frame buffer. The convergence angle was 4.7°, consistent with the 79 cm observation distance and an interpupillary separation of 65 mm. Care was taken in the optical alignment so that the small binocular probe, the only stereo feature presented, could be fused immediately.

Procedure: Four paid subjects participated in the following experiments. All had good stereo vision and were naive to the goals of the experiments. For these experiments the subjects made forced-choice judgments of whether the stereo depth probe was in front of or behind the stimulus surface. The surface was identical with that used in the slant experiment, presented in both orthographic and perspective projection. The depth probe was a binocularly-projected vertical bar subtending 18′ by 1′ (Fig. 3). Subjects saw the surface in *3D* quite vividly and several were unaware of the fact that all but one of the curves were actually viewed monocularly. When the bar was projected binocularly with a given stereo disparity it immediately appeared in depth either farther than or nearer to the surface, depending on the location at which the depth was probed.

The basic task was therefore to visualize the stimulus surface in space during the first interval of 500 ms, then when the probe appeared for an additional 750 ms, the subject was to compare the (stereoscopic) depth of the probe with the depth of the surface, and to response with mouse buttons whether the probe was farther or nearer. As before, a third button could be depressed if the subject was too uncertain of the relative depth judgment, whereupon the trial was not recorded. The depth judgments were made at four equally-spaced probe locations lying along a straight line (ruling) on the surface, parallel to the ridges and troughs of the sinusoidal cylinder, as shown in Fig. 3. The central two probe locations were approximately 100′ from the center of gaze; the outermost location was at an eccentricity of 140′ (perspective) and 180′ (orthographic).

Three depth experiments were performed. In the first experiment the stimulus surface was not purely

---

[2] The lack of any visible motion parallax, particularly for near observation distances, could in principle have served as a cue revealing that the probe and the stimulus surface were equidistant. Depth differences were successfully measured despite this potential cue

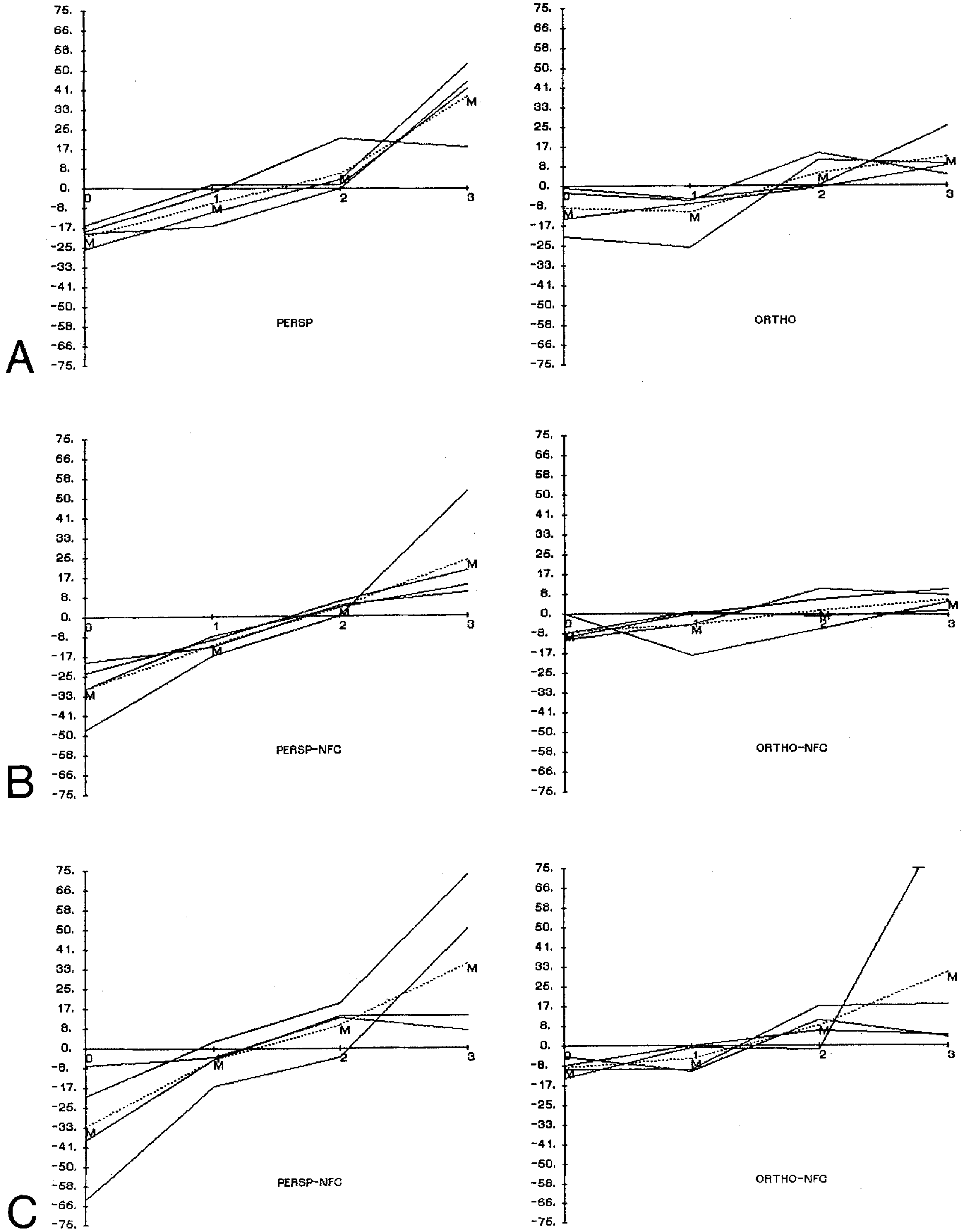


Fig. 4A–C. Graphs of mean disparity as a function of probe location 0–3 for perspective and orthographic projection. In **A** the stimulus on time was 500 ms with probe on time of 750 ms. In **B** the stimuli were presented without fixation contours. The third experiment, **C**, was like **B** but the stimulus and probe on times were reduced to 100 and 150 ms monocular; one of the contours, the central contour, straddled by the two middle probe locations, was projected binocularly with zero stereo disparity. This single contour served both as a fixation reference and, presumably, as a means to anchor the overall surface in depth. In the second experiment the fixation contour was removed and subjects were instructed to fixate at the center of the screen.

While in the first two experiments the stimulus was presented for 500 ms and the probe remained an extra 750 ms, in the third experiment the stimulus was presented for 100 ms followed by an additional 150 ms with the depth probe superimposed. This experiment was prompted by questions regarding the role of eye movements in the depth judgments. The 500–750 ms condition used in the first two depth experiments allowed subjects to shift their direction of gaze to comfortably fixate the stereo probe after it had appeared at a probe location at random. In the 100–150 ms condition, the 150 ms of probe exposure provided insufficient time to shift fixation from the center of the screen to the probe.

As in the slant experiment, the depth experiments used randomized-staircase forced-choice method. The apparent depth, quantified by stereo disparity, was measured at the various probe locations by 10 transitions between nearer-than and farther-than, and vice versa, at each of the four probe locations along the surface. Each of the four subjects was run on four repetitions of each of the experiments.

Results and Discussion

The orthographic and perspective data for each experiment were analyzed separately. For each subject, the transition disparities at each probe location were averaged across the four runs. The graphs in Fig. 4 show the results for the four subjects as separate curves, with the abscissas corresponding to the four probe locations (as in Fig. 3) and the ordinates corresponding to the mean stereo disparity, in arc minutes, at which the stereo probe appeared to intersect the monocular surface. Each plotted point is the mean of four runs, where the data for each run, at each location, was the mean of 10 transitions or reversals in direction. The dotted line labelled "*M*" in each graph is the combined mean across the four subjects.

The results of the first experiment, in which a zero-disparity stereo fixation contour was provided, is shown in Fig. 4a. An immediately observable effect is that in the perspective projection measured depth increased monotonically across the four probe locations, with probe locations 0 and 1 seen as nearer (as indicated by negative mean probe disparities). A large gradient of depth existed between locations 1 and 2 – the mean disparity difference of 17.7′ corresponded to an angle of 75° out of the image plane (which constitutes a greater angle, in fact, than the geometrically computed angle of 63.5° between the image plane and the line connecting points 1 and 2 on the surface). In the orthographic case, the measured depth gradient was similiarly large between the central locations, but there was generally little increase in apparent depth at eccentricity (between locations 2 and 3 on the right, and 0 and 1 on the left). In contrast with the perspective data, orthographic depth tended to be steep through the center but to flatten with eccentricity. A second observation is that the variability, both within subject and across subjects, tended to increase with eccentricity.

In the second depth experiment (Fig. 4b), the stimuli were purely monocular ("NFC" refers to "no fixation contour"). The results without the fixation contour were surprisingly better, overall, than in the first experiment. It is particularly noteworthy that the data crosses zero between probe locations 1 and 2, with negative disparities at locations to the left of the center of the monocular surface and positive disparites at locations to the right of center. The center point, where the subjects fixated, was implicitly regarded as equidistant with zero stereo disparity, i.e., the subjects apparently matched the depth of the stereo horopter with the fixation point of the monocular surface. The monocular surface appeared to vary in depth about that point.

The third depth experiment (Fig. 4c) repeated the NFC condition but with only 250 ms total observation time, 150 ms of which the depth probe was available for comparison with the monocular surface. The data showed increased variability, particularly at eccentrity (locations 0 and 3). The mean disparity graphs (labelled "*M*") again crossed zero between probe locations 1 and 2.

The steep gradient of perceived depth between probe locations 1 and 2 noted earlier occurred in the NFC experiments as well, with corresponding angles out of the image plane varying from 59.7° to 74.1°. Comparing the mean disparity gradient across the three experiments, in either perspective or orthographic conditions, the means were not significantly different. As is evident from the graphs, the data is similar in the central region, suggesting to us that the depth gradient seen in the very short exposure was substantially the same as in the longer exposures, and likewise the depth seen in the orthographic stimuli was substantially the same as in the perspective. The means across subjects are not very revealing, however, in light of the large subject differences we found. Figure 5 shows the data for the individual subjects, grouped vertically by experiment. For a given subject and task the data is generally quite systematic, but comparing

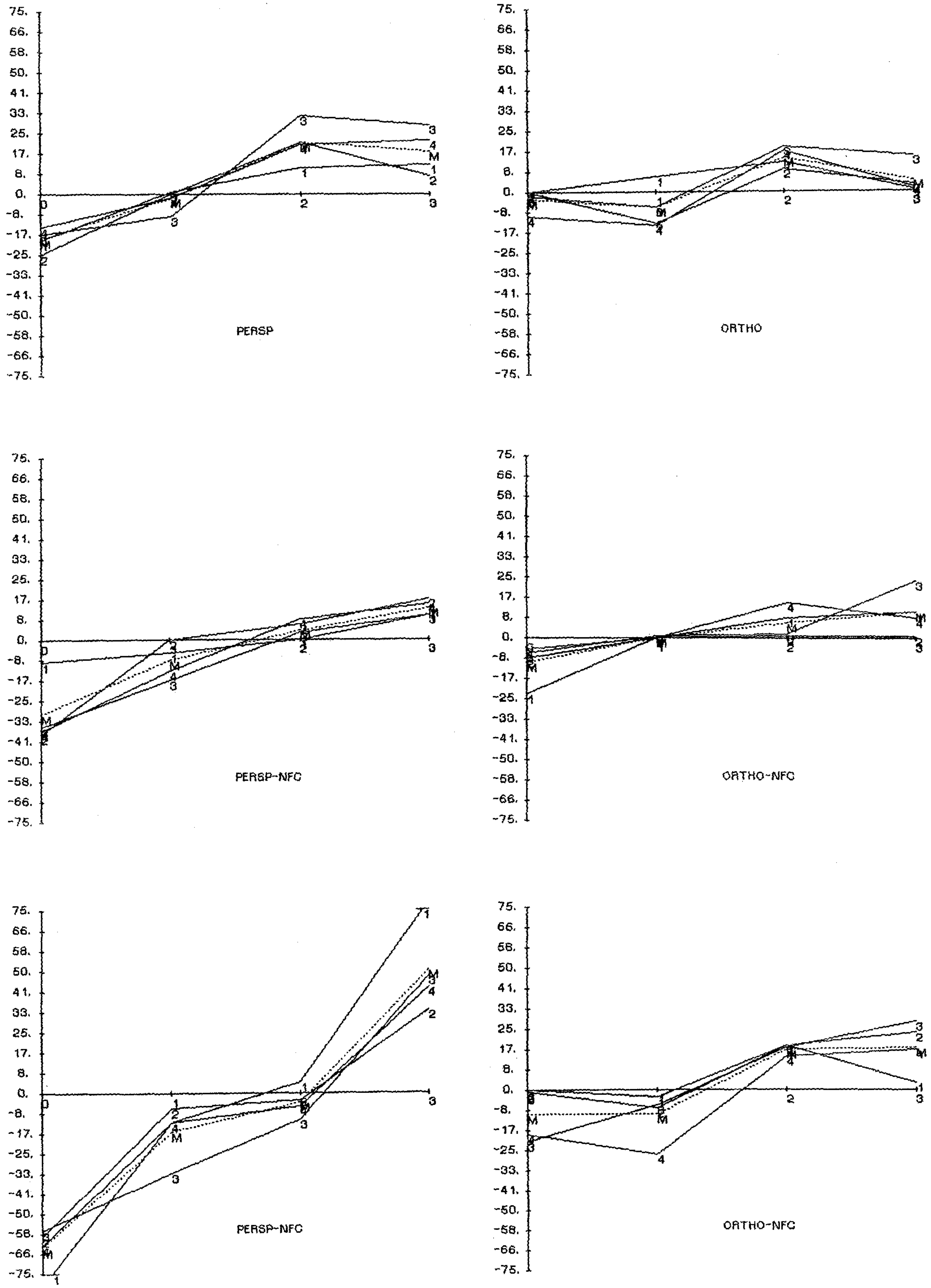
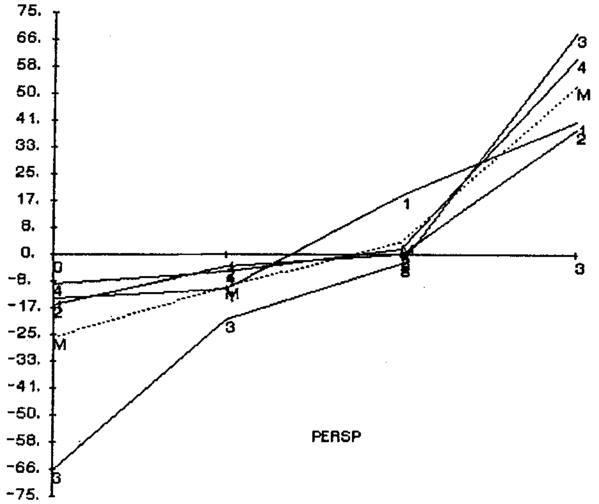
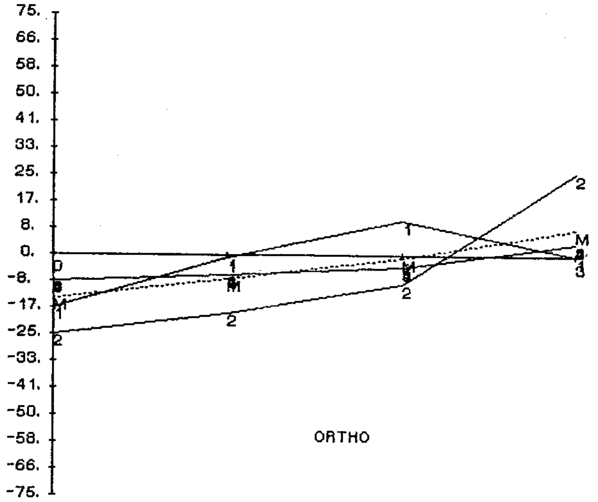
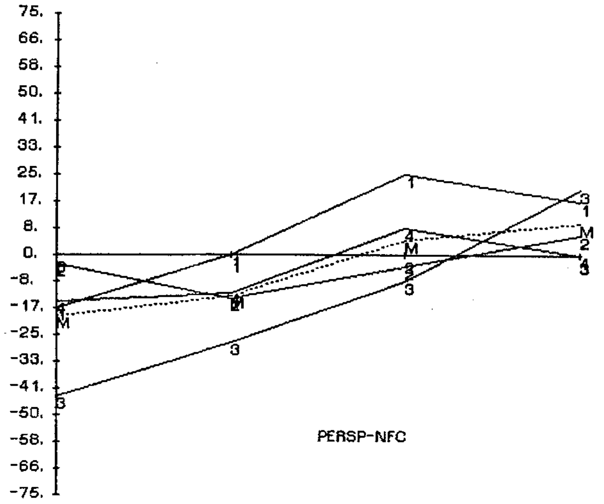
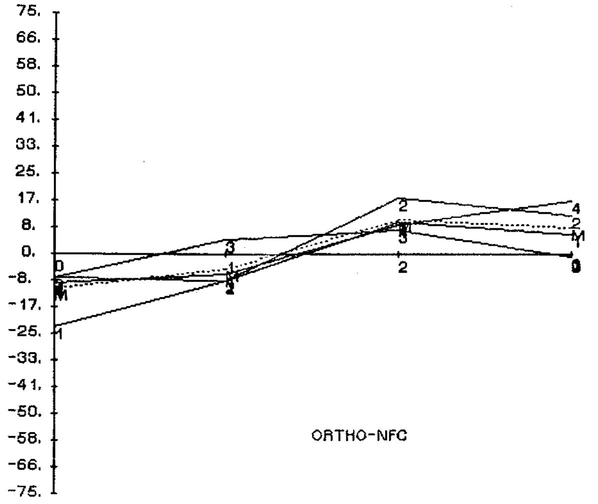


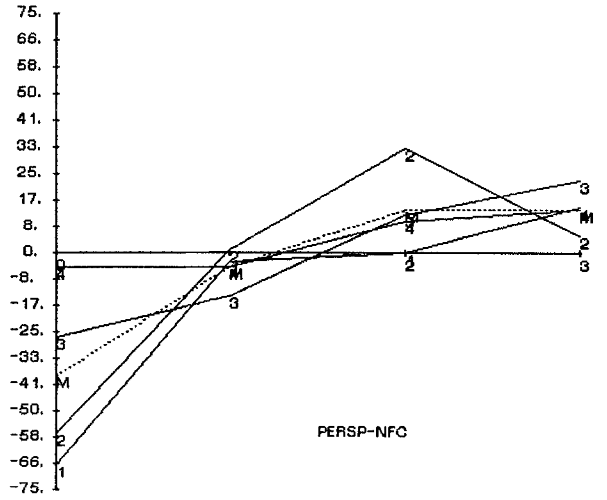
Fig. 5. Individual subject performance for the three depth experiments, arranged as in Fig. 4

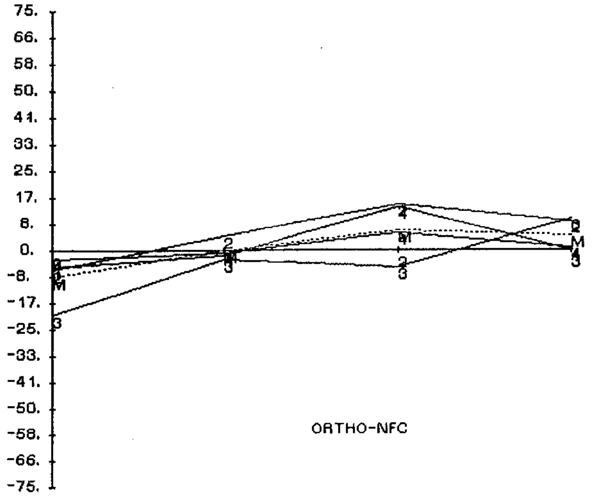

75.
66.
58.
50.
41.
33.
25.
17.
8.
0.
-8.
-17.
-25.
-33.
-41.
-50.
-58.
-66.
-75.
0
1
2
3
M
4
PERSP

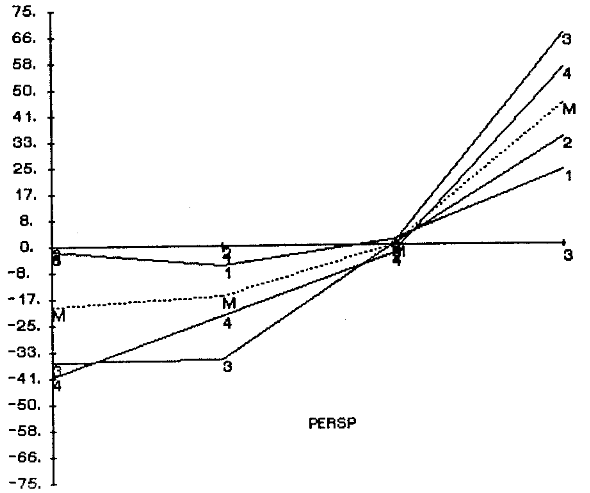

75.
66.
58.
50.
41.
33.
25.
17.
8.
0.
-8.
-17.
-25.
-33.
-41.
-50.
-58.
-66.
-75.
0
1
2
3
M
ORTHO

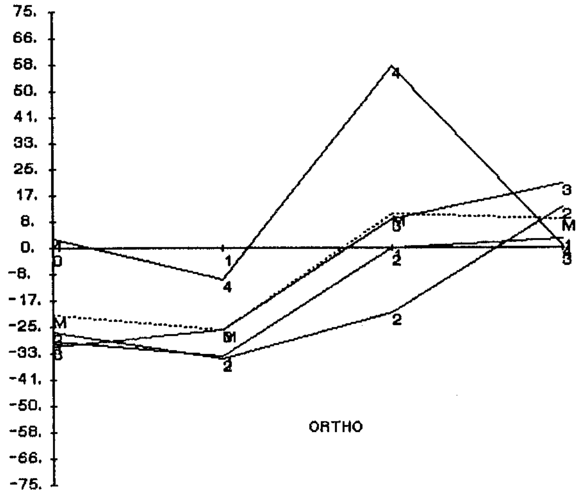

75.
66.
58.
50.
41.
33.
25.
17.
8.
0.
-8.
-17.
-25.
-33.
-41.
-50.
-58.
-66.
-75.
0
1
2
3
4
M
PERSP-NFC

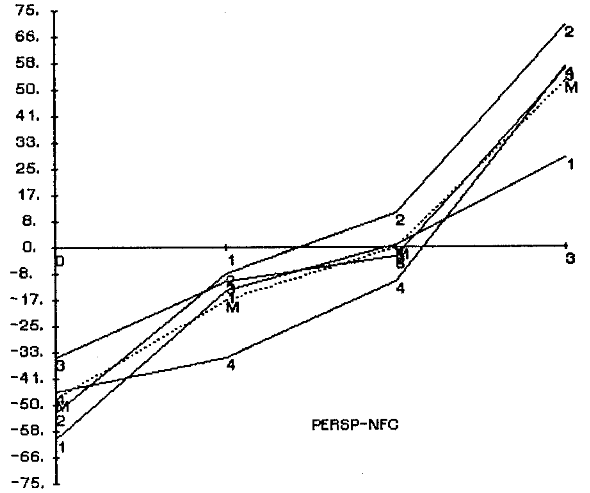

75.
66.
58.
50.
41.
33.
25.
17.
8.
0.
-8.
-17.
-25.
-33.
-41.
-50.
-58.
-66.
-75.
0
1
2
3
4
M
ORTHO-NFC

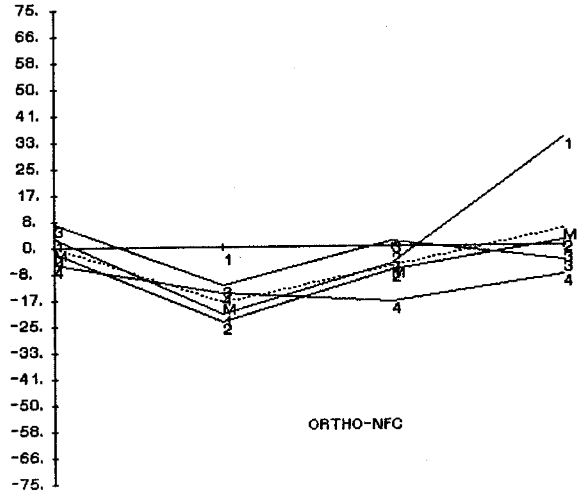

75.
66.
58.
50.
41.
33.
25.
17.
8.
0.
-8.
-17.
-25.
-33.
-41.
-50.
-58.
-66.
-75.
0
1
2
3
4
M
PERSP-NFC

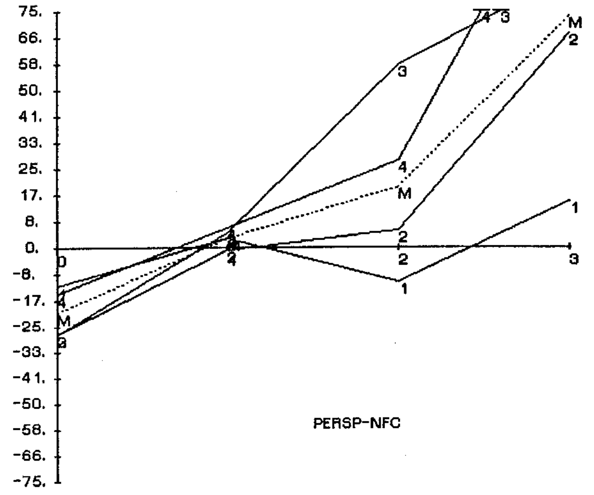

75.
66.
58.
50.
41.
33.
25.
17.
8.
0.
-8.
-17.
-25.
-33.
-41.
-50.
-58.
-66.
-75.
0
1
2
3
4
M
ORTHO-NFC

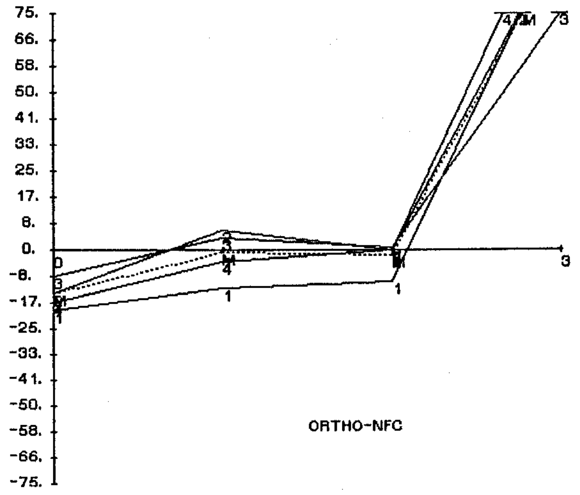

75.
66.
58.
50.
41.
33.
25.
17.
8.
0.
-8.
-17.
-25.
-33.
-41.
-50.
-58.
-66.
-75.
0
1
2
3
1
2
3
4
M
PERSP

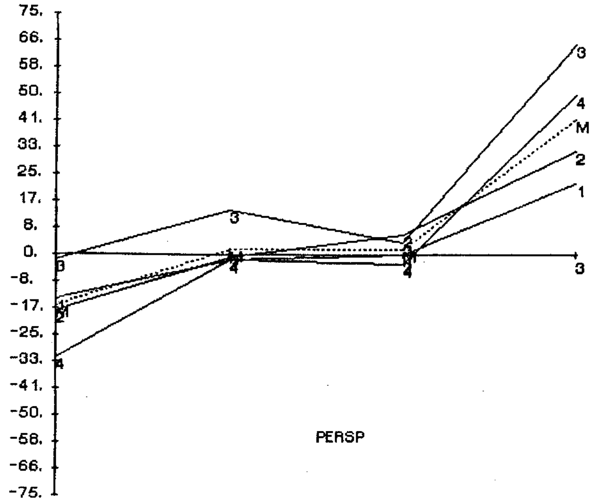

75.
66.
58.
50.
41.
33.
25.
17.
8.
0.
-8.
-17.
-25.
-33.
-41.
-50.
-58.
-66.
-75.
0
1
2
3
1
2
3
4
M
ORTHO


75.
66.
58.
50.
41.
33.
25.
17.
8.
0.
-8.
-17.
-25.
-33.
-41.
-50.
-58.
-66.
-75.
0
1
2
3
1
2
3
4
M
PERSP-NFC

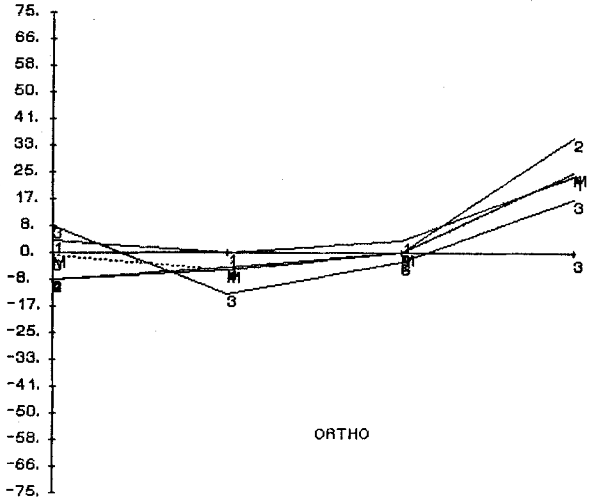

75.
66.
58.
50.
41.
33.
25.
17.
8.
0.
-8.
-17.
-25.
-33.
-41.
-50.
-58.
-66.
-75.
0
1
2
3
1
2
3
4
M
ORTHO-NFC


75.
66.
58.
50.
41.
33.
25.
17.
8.
0.
-8.
-17.
-25.
-33.
-41.
-50.
-58.
-66.
-75.
0
1
2
3
1
2
3
4
M
PERSP-NFC

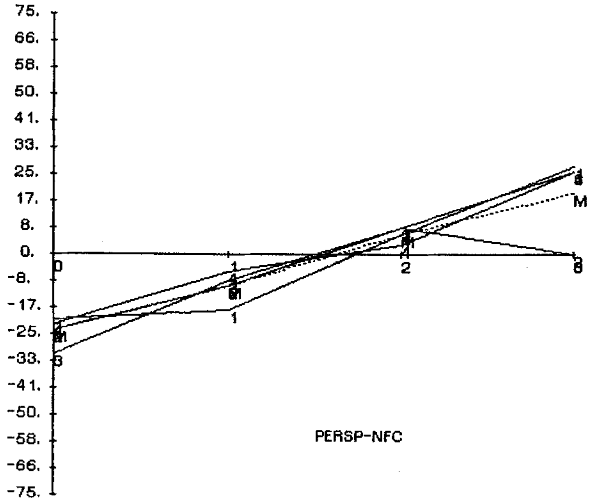

75.
66.
58.
50.
41.
33.
25.
17.
8.
0.
-8.
-17.
-25.
-33.
-41.
-50.
-58.
-66.
-75.
0
1
2
3
1
2
3
4
M
ORTHO-NFC



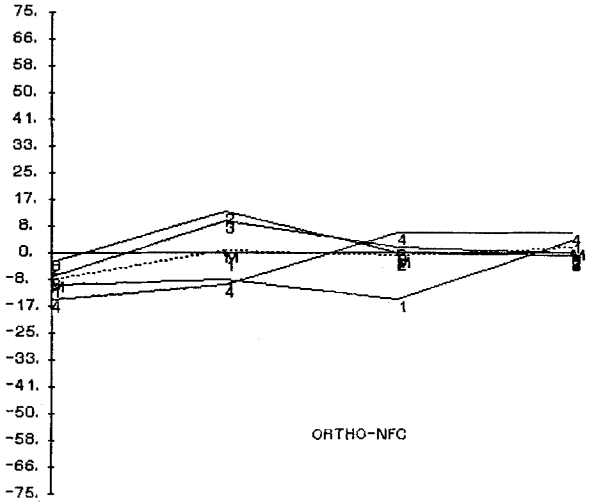
PERSP



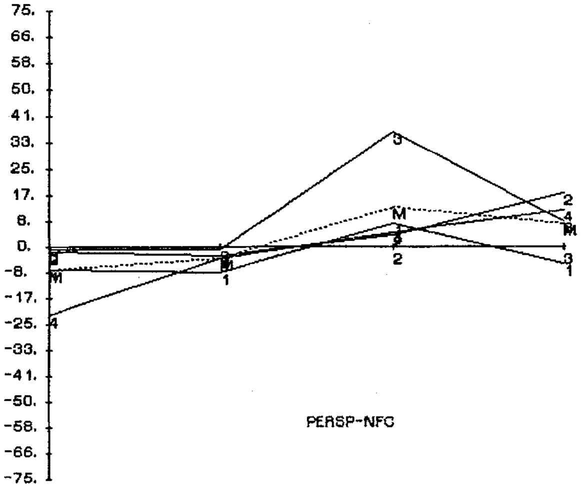
ORTHO



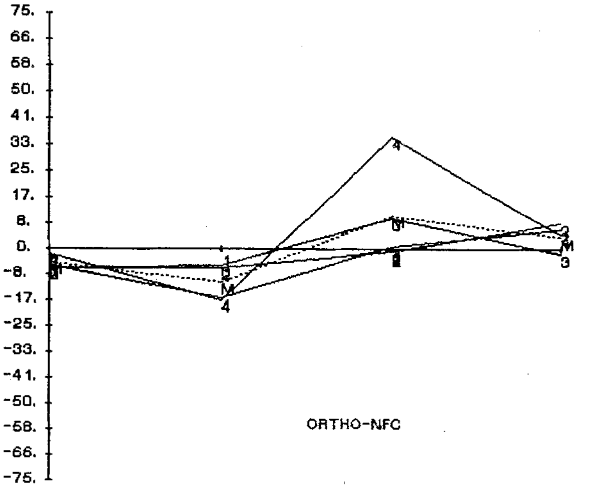
PERSP-NFC



ORTHO-NFC



PERSP-NFC



ORTHO-NFC

across subjects the data were often quite different. Some of the differences could reflect artifacts of not controlling for eye movements, despite the experimental instructions to attempt to maintain central fixation. We are intrigued, however, with the robust tendency for perspective depth to increase with eccentricity while orthographic depth tends to first increase then either reach a plateau or decrease. Assuming that eye movements were similar for these two conditions, the results suggest a different perception of depth in the periphery of the orthographic images.

2.3 General Discussion

Figure 1a is a perspective projection of a smooth surface that seems to extend in depth. Figure 1b, the equivalent orthographic projection, appears very similar in 3-*D* shape, but without the progressive decrease in scale with distance seen in Fig. 1a the surface does not appear to extend in depth as convincingly. This informal impression was reflected in the results of the depth experiment. While the orthographic image might not vary as dramatically in overall depth, in any region the surface appears strongly foreshortened. The slant experiment showed that the perceived slant in these images, both orthographic and perspective, is quite substantial and, modulo the precision of the experiment, an accurate reflection of the geometrically-correct slant. Apparent slant in any vicinity of the surface is a measure of the gradient of depth, i.e. the rate of change of depth (and being the derivative of distance, slant is independent of the absolute distance to the surface). Thus, in perceiving the slant of a given surface patch in either orthographic or perspective projection, one has an appreciation for the local direction and rate at which depth increases – which might be considerable if the surface appears strongly slanted. But while the local gradient of depth might be considerable, the surface might not appear to accumulate a large overall depth difference between two such localities – paradoxically slanted but without depth. That paradox is resolved by hypothesizing that the visual system has two fundamentally distinct perceptual activities, the appreciation of local surface orientation and the appreciation of depth differences, and that the two are not necessarily maintained in perfect register or correspondence. Thus while surface orientation and depth covary geometrically, to accumulate a depth difference (and not just a measure of the local gradient of depth) requires line integration of the depth gradient along a path connecting the two points in question. The computation of a depth difference from slant might well be imprecise compared to the perceived local surface orientation. This was found in comparing the apparent slant and apparent depth results, particularly for the orthographic stimuli.

Consider what is required to perform the depth comparison task in the case of the purely monocular stimuli. The straightforward solution would be to localize the monocular surface in absolute distance from the observer, to likewise localize the stereo probe in absolute distance, then to compare the two absolute distances for the given probe location. For distances up to approximately 2 m, stereopsis provides reasonably precise perception of absolute distances (Wallach and Zuckerman 1963; Ritter 1977, 1979; Morrison and Whiteside 1984) as well as absolute distance increments (Ono and Comerford 1977; Wallach et al. 1979). It is therefore reasonable to expect that the stereo probe is accurately localized in absolute depth, measured either from the observer or as a signed absolute distance interval relative to the absolute distance of the horopter. The problem is in localizing the monocular surface in absolute depth so that it might be compared. This problem arises in both the orthographic and perspective projection[3]. The fact that there is insufficient information to perform the depth matching task, and the experimental observation that we evidently can, suggests that the monocular percept is fixed in the same space as the stereo probe. The assumption used to link these two spaces, apparently, is that the absolute distance of the monocular surface at the fixation point equals that of the stereoscopic horopter at the same point. This is evident in the graphs of disparity versus probe location, where the zero intercept of disparity occurred close to midpoint between probe locations 1 and 2, which corresponded to the central fixation point of the monocular image. This hypothesis seems sound in that whatever surface location is fixated in sharp focus is likely to lie at zero disparity, since in the near field at least, there is close coupling between vergence and accommodation that brings into sharp focus the fixation point that is also at zero disparity. The fixated surface point (seen monocularly in our stimuli but binocularly in normal vision) is thus assumed to be at the absolute distance of the horopter. With the two depth measures (stereo and mono) sharing a common zero intercept, it is then a matter of calibrating the proportionality between monocular depth and disparity.

By this scheme, the forced-choice judgment of whether the probe is nearer or farther than the surface

---

[3] Note that while static perspective projection provides information about ratios of absolute distances, absolute distance information is generally not available. The exception is when an object of known absolute size is projected, whereupon the absolute distance can be recovered from the ratio of known size/retinal size

reduces to comparing two absolute depth increments. The stereo disparity of the probe is converted to a signed absolute distance interval, with the zero value corresponding to the horopter, and likewise the monocular depth is treated as a signed relative distance interval at the corresponding surface location, with the zero value corresponding to the center fixation point of the surface (in the above experiments). In the depth experiments we found that apparent depth increased quite considerably with displacement across the surface.

We draw three general conclusions from these experiments. First, depth is derived from orthographic projection as a scaled quantity that is commensurate with the depth perceived from stereopsis in the near field. Second, the comparison of monocular and stereo depth is rather fast (achievable with exposures of only 150 ms)[4] and does not require eye movements. The third conclusion is that the absolute distance to a fixated monocular surface is assumed to coincide with the stereo horopter. Binocular vision generally puts a fixated surface point in sharp focus and at zero disparity. Likewise, a fixated surface point in a monocular image, seen in sharp focus, is apparently regarded as lying at the same absolute distance as it would be if viewed binocularly at zero disparity. We regard these conclusions as quite tentative. The role of the fixation point in the depth comparison task is not clear, and merits a study that carefully monitors eye movements. We found, in a pilot experiment, that as subjects fixated a given probe location, instead of the center of the surface, the probe was more likely to appear to lie at the same depth as the monocular surface (perhaps related to Emmert's Law).

It is also not clear why apparent depth in orthographic projections first increased then decreased with eccentricity for some subjects and yet increased monotonically for others. This behavior was not apparent in the slant experiment. Presumably the strategy for performing the depth comparison task was the same regardless of whether the projection was orthographic or perspective.

---

[4] It has been long recognized that very brief binocular presentations can result in an impression of stereo depth. We note here that the impression of depth from stereo disparity can also be compared with that derived monocularly in very brief presentations

References

Chevrier J, Delorme A (1983) Depth perception in Pandora's box and size illusion: evolution and age. Perception 12:177–185

Deregowski J (1970) Hudson's pictures in Pandora's box. J Cross-Cultural Psychol 1:315–323

Gregory R (1968) Visual illusions. Sci Am 218 (11):66–76

Gregory R (1970) The intelligent eye. Weidenfeld and Nicolson, London

Kennedy JM (1974) A psychology of picture perception. Jossey-Bass, San Francisco

Morrison JD, Whiteside TCD (1984) Binocular cues in the perception of distance of a point source of light. Perception 13:555–566

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

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. Percep Psychophys 25:209–214

Stevens KA (1981) The visual interpretation of surface contours. Art Intell 217 (Special Issue on Computer Vision):47–74

Stevens KA (1983a) Slant-tilt: the visual encoding of surface orientation. Biol Cybern 46:183–195

Stevens KA (1983b) The line of curvature constraint and the interpretation of 3-*D* shape from parallel surface contours. Eighth International Joint Conference on Artificial Intelligence, August, pp 1057–1061

Stevens KA (1983c) Tilt (the direction of surface slant): a neglected psychophysical variable. Percep Psychophys 33:241–250

Stevens KA (1986) Inferring shape from contours across surfaces. In: Pentland AP (ed) From pixels to predicates: recent advances in computational vision. Ablex, Norwood, pp 93–110

Topper DR, Simpson WA (1981) Depth perception in linear and inverse perspective pictures. Perception 10:305–312

Wallach H, Gillam B, Cardillo L (1979) Some consequence of stereoscopic depth constancy. Percep Psychophys 26:235–240

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

Received: October 27, 1986

Prof. Dr. Kent A. Stevens
Department of Computer Science
University of Oregon
Eugene, OR 97403
USA