Pictorial depth cues: a new slant

G. Lee Zimmerman

Department of Electrical Engineering, Tulane University, New Orleans, Louisiana 70115

Gordon E. Legge

Minnesota Laboratory for Low Vision Research, Department of Psychology, University of Minnesota, Minneapolis, Minnesota 55455

Patrick Cavanagh

Vision Sciences Laboratory, Department of Psychology, Harvard University, Cambridge, Massachusetts 02138

Received August 19, 1993; revised manuscript received May 31, 1994; accepted June 1, 1994

Pictorial depth cues such as perspective projection, aspect ratio, and texture gradients can specify mathematically the slant of a planar surface. We performed experiments to measure the accuracy of human perception of surface slant from these cues. We calculated the perceived slant from judgments of the relative lengths of a pair of orthogonal lines embedded in the surface. Our results indicate that slant judgments are accurate to within 3 deg. This level of accuracy was achieved whether the cues were luminance differences or equiluminous color differences. We found no evidence of the recession to the frontal plane that has been reported by Gibson [J. J. Gibson, *The Perception of the Visual World* (Houghton Mifflin, Boston, Mass., 1950)] and others. We did find evidence suggesting that subjects do not make accurate depth estimates of disconnected surfaces. This may be the source of the discrepancy between our measures and those of Gibson and others. This research, combined with previous findings, supports a model of perception that involves at least two and possibly more representations of space: one local veridical representation of surface orientation derived primarily from pictorial cues and another global representation of observer-centered distance derived primarily from binocular disparity and motion parallax.

1. INTRODUCTION

The image projected on the retina is rich in information about the spatial structure of the surrounding world. This information is carried by a variety of depth cues. Although the information from a depth cue may specify mathematically the orientation of the surface being viewed, human perception might not make use of this information or may use it inaccurately. For example, binocular vergence and disparity can specify the distance of a point from the observer. However, Foley[1] found that perception of space from binocular disparity alone may incorporate non-Euclidean distortions. Perceived space may be different from physical space. To understand vision, we must understand the perceptual representation of space and its influence on visual performance.

In this paper we are concerned with the perception of slant derived from static monocular depth cues called pictorial cues. These are the depth cues that can be captured in a single photograph. Some examples are occlusion, shading, shadows, perspective distortion, texture gradients, and aspect ratios. An image of a natural scene has a great deal of information about the structure and orientation of surfaces without the observer's having to call on nonpictorial cues such as binocular disparity and motion parallax.

How accurate are observers at exploiting these cues? In 1950 Gibson[2] had subjects rotate a palm board to match the slant of a textured surface projected on a screen. He observed that the palm board settings indicated a perceived slant that was consistently closer to the frontal plane than to the optically specified slant. This underestimate was of the order of 10 deg and was confirmed by many other researchers.[3–5] Gibson called this phenomenon "recession to the frontal plane." He suggested that the recession was due to residual information for the frontal position of the screen. He was unable to remove this residual information.

Perrone[6] proposed an alternative explanation of recession to the frontal plane. He suggested that the observers were making errors with respect to the direction of the line of sight to the projected planes. The error would make the angle between the true line of sight and the plane larger than the angle between the perceived line of sight and the plane, as shown in Fig. 1. Reference errors such as these have also been reported by Matin and Li.[7] The line-of-sight error combined with the geometry of the tasks would account for the reduction of the perceived slant. This account is not completely satisfying because it does not explain why the misinterpretation should occur at all. Perrone's work also implies that if the perceived line of sight were correct, the perceived slant would be veridical.

Todd and Reichel[8] argued that the recession to the frontal plane is due not to line-of-sight reference errors or residual information from conflicting surface cues but to the nature of the representation of three-dimensional (3-D) space. They noted that slant judgment tasks are difficult to perform and lead to large angular deviations that are systematic underestimates of the surface slant.

0740-3232/95/010017-10$06.00

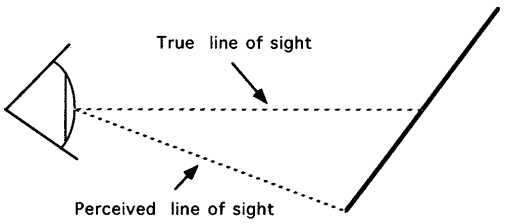

Fig. 1. Misperceived line of sight.

This led them to conclude that local mappings of depth or orientation are not the primary form of representation. Instead they suggest that visual perception uses an ordinal representation of space, which maintains relative depth ordering of surfaces while not representing the metric changes between these surfaces.

On the other hand, Stevens and Brookes[9] argued that accurate metric judgments could be made if the appropriate probe were used. They required subjects to make aspect-ratio judgments on projected circles painted on undulating surfaces, for which the surface slant was indicated through contour lines. In contrast to the results of Gibson[2] and others, Stevens and Brookes[9] found that their subjects were quite accurate and inferred that subjects maintain a precise perceptual representation of local surface slant. These local measurements of perceived surface orientation were compared with surface orientation calculated from binocular disparity measurements of the same surface. Surface orientation gathered from disparity measurements did not correspond to monocular measurements, strongly suggesting at least two separate representations of the spatial structure of a single surface.

This paper concentrates on the representation of surface orientation based on pictorial cues. If recession to the frontal plane is characteristic of depth representations built from pictorial cues, there will be a substantial effect on visually guided behavior. Is there recession to the frontal plane in perceptual slant estimation, and what accounts for the discrepancy between the findings of Gibson and those of Stevens and Brookes? Is slant representation quite precise as suggested by Stevens and Brookes or rather crude as might be expected from the kind of nonmetric coding suggested by Todd and Reichel? Is there evidence that slant judgments based on pictorial cues are distinct from depth based on other cues, as suggested by Stevens and Brookes? In this paper we address these issues and develop support for the concept of multiple independent representations of visual space.

2. EXPERIMENT 1: ESTIMATING SLANT BY COMPARING A TEST PLANE WITH A STANDARD PLANE

To determine the perceived slant of a surface from only static monocular cues, we had observers make judgments about the relative lengths of lines painted on two planar surfaces: a test plane with an unknown slant angle and an upright standard plane. The judgments were used to calculate the perceived slant angle of the test plane. The slant of the surfaces was depicted by texture and boundary cues. From previous studies[2–5,8] we expected that line-length judgments made parallel to the direction of the slant would show a marked difference as a result of the phenomenon of recession to the frontal plane. We also expected to find no such differences for line-length judgments made perpendicular to the direction of the slant of the surface.

A. Method

Two 3-D planar surfaces were synthesized graphically on a CRT screen. The 3-D planes were square, measuring 6 cm on a side, and were presented side by side on the screen, separated by 3 cm. Each subject viewed the planes monocularly, with the eye positioned between the two planes at a distance of 28 cm from the screen. In all the experiments presented in this paper the lines, the texture, and the surface were first drawn frontally and then slanted appropriately in space, so that all the texture gradients, aspect ratios, and angles in the image were appropriate for a surface viewed from the simulated distance and slanted at the appropriate displayed angle as shown. Each plane was rotated graphically 12 deg, as shown in Fig. 2, such that the surface of the plane was perpendicular to the line of sight and was then projected appropriately onto the screen surface. Each subject wore an eyepatch, and the subject's head was held in the appropriate position by a chin rest. All observations took place in a dimly lit room.

The standard plane was frontal (0-deg slant) to the subjects at all times and had a light-gray line of 3-cm length painted graphically on its surface. The test plane had painted on its surface a line that could vary in length. The test plane was slanted at 0, 30, or 60 deg about its horizontal axis. The standard plane was always viewed frontally. It subtended 9 deg on a side, and the line subtended 4.5 deg. The lines were light gray (73.2 cd/m^2), the texture elements were gray (25.7 cd/m^2), the plane was dark gray (5.49 cd/m^2), and the background was black (0.519 cd/m^2). These luminances were chosen so that all the elements could be clearly distinguished.

The slants of the planes were depicted graphically through texture cues and boundary cues. The texture cues were randomly placed gray squares (0.5 cm × 0.5 cm) that covered 5% of the surface. The image of the plane was projected onto the CRT screen so that the texture elements exhibited the same compression, foreshortening, and density as would a real slanted plane viewed from the specified point of view. The boundary cues were created by the abrupt decrease in intensity between the edge of the planar surface and the background.

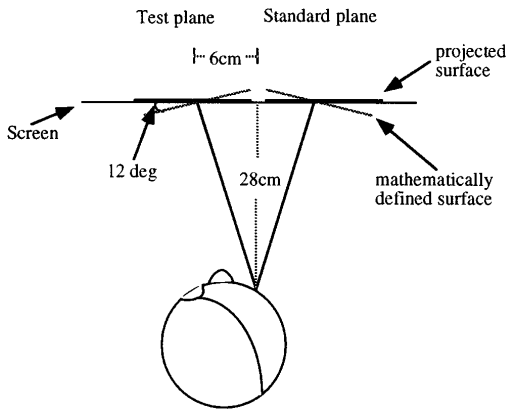

Fig. 2. Top view of the experimental setup.

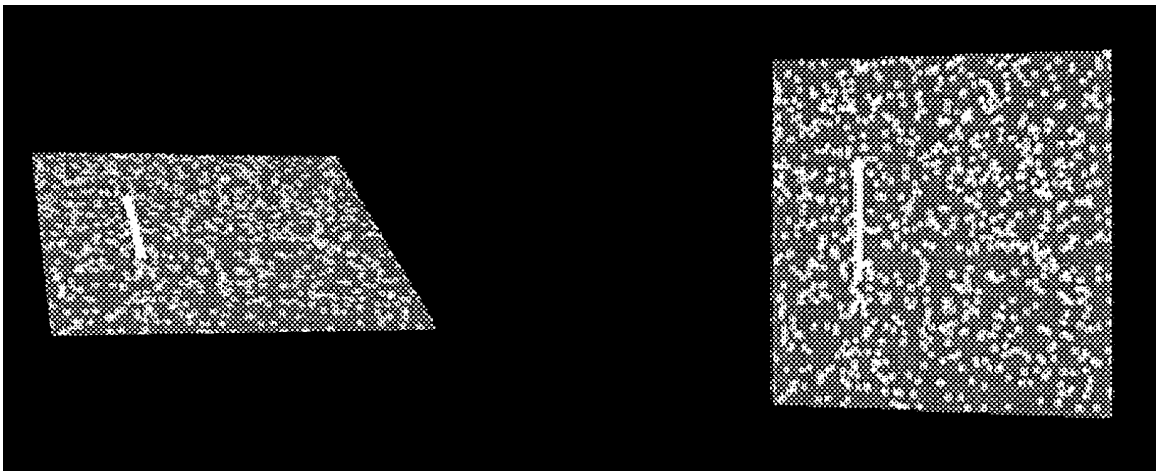

Fig. 3. Example of the stimulus used for experiment 1.

The boundary of the plane contained both perspective cues and foreshortening cues that indicted the slant of the surface. An example of the stimulus is shown in Fig. 3.

The probe lines were 0.05 cm wide and were light gray, making them easily visible on the textured surfaces. A vertical line (i.e., parallel to the tilt of the plane) on the standard plane was matched to a vertical line on the test plane. A horizontal line (i.e., perpendicular to the tilt of the plane) on the standard plane was matched to a horizontal line on the test plane. The line on each plane was shifted from the center of the plane in a random direction and radius not exceeding 1 cm on each trial.

Subjects were instructed to compare the length of the line on the test plane with the length of the line on the standard plane as if the lines were painted on the 3-D surface of each plane. The observers indicated whether the test line was longer or shorter than the standard line in a 3-D sense.

In experiment 1 we used the method of constant stimuli in a two-alternative–forced-choice paradigm. On each trial, one of seven lines of various lengths was selected at random and painted onto the test plane. The subject reported whether the test probe was longer or shorter than the 3-cm line painted on the standard plane. The set of seven lengths was chosen from pilot testing so that they spanned the point of subjective equality. A psychometric function, percent judged longer than the standard as a function of test-line line length, was based on 100 trials. This was done for each condition and slant angle. Trials at different slant angles were randomly interleaved within a single session. The psychometric functions were fitted by a cumulative normal function. A match was defined as the 50% point of the fitted curve.

These matched lengths were used to calculate the perceived surface slant. The calculations and their geometric derivations are shown in Fig. 4 and in Eqs. (1) and (2). In Fig. 4, La is the 3-D matched length set by the subject, θa is the actual slant of the plane, da is the true distance from the observer's eye to the test plane, Lp is the perceived length of the line on the test plane that is assumed to be equal to the length of the line on the standard plane, θp is the perceived slant angle of the plane, and dp is the perceived distance from the observer's eye to the test plane. We know La, da, and θa from the geometry of the plane and the lines. From this knowledge we can derive the angle δ subtended by the displayed line, using Eq. (1). Equation (2) calculates the perceived slant angle of the plane, θp, from the angle δ, the perceived line length Lp, and the perceived distance to the plane dp. In experiment 1 we assume that the perceived distance to the plane, dp, is equal to the actual distance to the plane, da. This assumption may not be true (we discuss evidence concerning this assumption below). In Eqs. (1) and (2) the trigonometric functions take values in degrees rather than in radians.

$$\delta = \tan^{-1}\left[\frac{La\,\cos(\theta a)}{da + La\,\sin(\theta a)}\right], \tag{1}$$

$$\theta p = 90 - \delta - \sin^{-1}[dp\,\sin(\delta)/Lp]. \tag{2}$$

For example, assume that the test plane was slanted 60 deg ($\theta a = 60$ deg) at a distance of 28 cm from the subject ($da = 28$ cm) and that the subject matched a 4-cm line on the test plane ($La = 4$ cm) to a 3-cm line on the standard plane ($Lp = 3$ cm). From Eq. (1) the angle δ subtended by the matched line is 3.64 deg. We assume that dp is equal to 28 cm, the actual distance to the plane. From Eq. (2) the calculated perceived slant angle θp is 50.06 deg, a significant recession to the frontal plane.

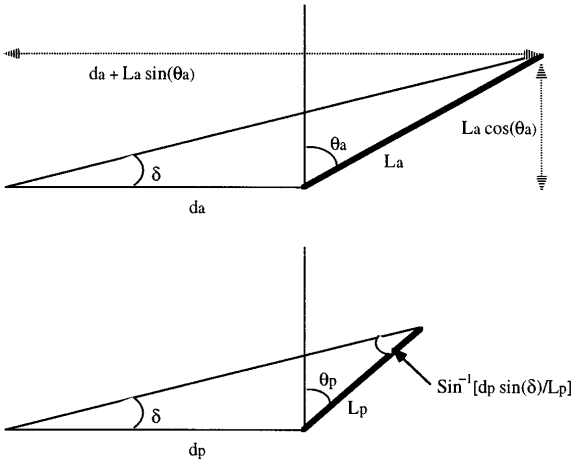

Fig. 4. Geometry of surface-slant estimates from match lengths.

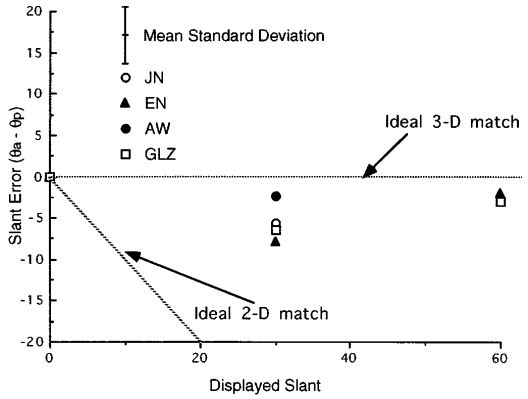

Fig. 5. Slant error, defined as perceived slant, θa, minus displayed slant, θp, derived from the vertical–vertical matches, for four observers.

Four subjects participated in these measurements. Three were naïve to the goals of the study. All had normal or corrected-to-normal vision. All the subjects were asked to treat the planes and lines in the display as if they were actual 3-D objects. None of the subjects reported any difficulty doing the task.

B. Results

We considered two situations, one in which the test line and the standard line were vertical and one in which the test line and the standard line were horizontal. If subjects underestimate the slant of the test plane, the two-dimensional (2-D) projection of the test line at match will be longer than the projection of the line with veridical match. An increase in matched line length translates into an underestimate of slant, i.e., a recession to the frontal plane. This is exactly what we found. Figure 5 is a graph of the slant error (perceived slant minus displayed slant). The data in Fig. 5 show a marked recession to the frontal plane. The underestimate averaged ~7 deg for the four subjects. There were no systematic differences in the standard deviations of the judgments. The average standard deviation was 3.35 deg. Although these results show evidence of recession to the frontal plane, subjects clearly were using information besides the 2-D line lengths. If they had been matching only on the 2-D projections, the perceived slant angles would have been zero and the data would have fallen on a line with a negative 45-deg slope from 0-deg slant.

When the test and the standard lines were horizontal, the slant of the test plane did not alter the length of the test line in the image. We would expect that the horizontal-line judgments would be unaffected by slant. However, this was not the case, as shown in Fig. 6. At the point of subjective equality, the screen projections of the test line averaged 6% longer than the standard line for 30-deg slant and 10% longer for 60-deg slant. This is surprising and directly challenges some of our assumptions about the planes—specifically, our assumption that the test plane and the standard plane are at the same perceived distance from the observer. Subjects reported seeing the slanted test plane as being farther from them than the standard plane. This difference in absolute depth with slant angle could be due to a variety of depth cues such as height in the picture plane (higher on average for the slanted test plane) and absolute size in the image (smaller for the slanted test plane).

C. Discussion

The results of experiment 1 cast doubt on our ability to measure the perception of slant by using a comparison of targets on separated surfaces. The horizontal-match data strongly suggested that the test plane was perceived as being at a distance different from that of the standard plane. The computation of perceived slant in Eq. (2) depends on the perceived distance to the test plane:

$$\theta p = 90 - \sin^{-1}\left[\frac{dpLa}{daLp}\cos(\theta a)\right]. \tag{3}$$

If we assume that the true distance to the plane is much greater than the length of the projected line (i.e., the angle δ subtended by the probe is small) then Eq. (2) will approximate Eq. (3). When we examine Eq. (3) we can see clearly that an overestimate of the perceived distance to the test plane, dp, will produce recession to the frontal plane.

Methods for measuring perceived surface slant that involve the comparison of one surface with another rely on veridical distance estimates. If distance estimates are not veridical (or if one plane is perceived at a distance different from that of another plane), slant estimates will be inaccurate as well. However, the inaccuracy in slant estimation may be due not to an inaccuracy in local representation of surface orientation but instead to the inaccuracy in absolute depth estimates.

The perceived change in distance from the horizontal matched data is significant and interferes with our measurement of perceived slant. We are left with three possibilities: (1) the visual system may not be very good at encoding information about surface slant from pictorial cues, (2) the information may be encoded with considerable precision but with a frontal bias, or (3) local slant information may be represented accurately given the absolute depth estimate, but the absolute depth estimates from pictorial cues may be inaccurate.

3. EXPERIMENT 2: ESTIMATING SLANT FROM PROBE STIMULI ON A SINGLE PLANE

The changes in the perceived depth between the adjacent surfaces found in the first experiment will confound an

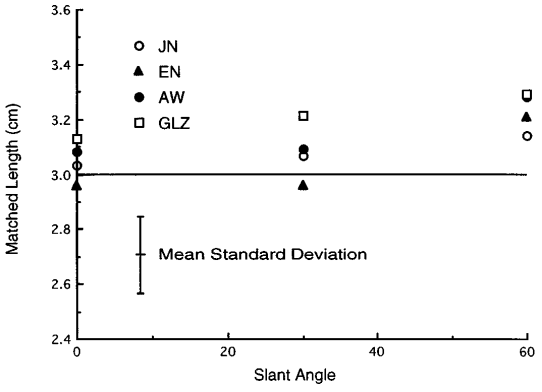

Fig. 6. Horizontal–horizontal matches, for four observers.

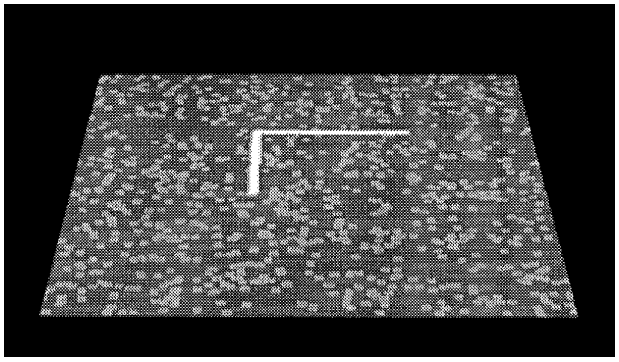

Fig. 7. Example of stimulus used in the single-plane experiment (experiment 2) with a Γ-shaped test probe.

accurate measurement of slant perception and will complicate any measure in which subjects are asked to compare two independent surfaces. To measure perceived slant we must either account for perceived distance changes or create a task that avoids distortions resulting from these changes. In our second experiment we used a single-plane method in which slant estimates could be computed independently of absolute-distance estimates. Because the probes are on the same slanted surface, changes in the perceived distance will affect the probe and the plane equally, providing the appropriate measure of perceived slant.

A. Method

The stimuli, the analysis, and the task were almost identical to those used in the first experiment. The subjects were presented with a single square plane in the center of the screen that was identical to the test plane of the first experiment. Instead of a line, a Γ-shaped test probe was graphically painted on the surface of the plane. The horizontal arm of the Γ maintained a constant length throughout the trials. The vertical arm of the Γ varied in length. The test probe was moved to a random position near the center of the plane on each trial. An example of the stimulus used in the second experiment is shown in Fig. 7. To calculate the perceived slant of the plane, we needed to find the length of the vertical line that appeared to be equal to the length of the horizontal arm of the test probe. From this value we could calculate the perceived slant angle of the plane.

The two arms of the probe were on the surface of the same plane, and the arms touched at a single point on the plane. This reinforced the perception that the horizontal arm and the vertical arm were at the same distance away from the observer on the slanted surface. If the two distances are equal, their exact value should be inconsequential to our estimate of slant as long as the angle subtended by the probe is small, as shown in Eq. (3). For this experiment we used the actual distance of the subject from the screen for da and dp in Eqs. (1) and (2) to derive the perceived slant θp.

For a plane with zero slant, observers matched a slightly shorter vertical line to the horizontal line. For our conditions the average vertical match was 87% of the horizontal line length. This is an example of a horizontal–vertical illusion. We compensated for this compression by using the measurement of the vertical line on the frontal plane (0-deg slant) as the actual length, Lp, for all of the calculated perceived slants of the individual subjects. We found the point of subjective equality by using two different methods, which produced similar results.

With the first method, observers were asked to judge whether the vertical line was longer or shorter than the horizontal line along the surface of the plane. This measurement was made with a two-alternative–forced-choice paradigm. The psychometric function was created (as it was in the first experiment) by the method of constant stimuli. The matched-length point was taken to be the 50% value of the psychometric function. All the other parameters for the experiment were identical to those used in the first experiment.

With the second method, subjects were able to change the length of the vertical line dynamically by moving a mouse. They used the method of adjustment to set the point of subjective equality. The subjects changed the length of the vertical line until they perceived it as being equal to the horizontal line. For this set of experiments the viewing distance was 26 cm, the plane was 11 cm square, and the texture consisted of 0.1 cm × 0.1 cm squares that randomly covered 20% of the area of the plane. The results from both methods are shown in Fig. 8.

B. Results and Discussion

The slant-error results for experiment 2, obtained with both the method of constant stimuli and the method of adjustment, are plotted in Fig. 8. The slant estimates made by all the subjects were within 2 deg of the actual slant of the plane with both methods. The standard deviation is 3.3 deg for all subjects and conditions of this experiment. There was no significant change in the standard deviation of the judgments as a function of slant angle.

The accuracy of the perceived slant is surprising, given most of the previous literature on perceived slant judgments, but it agrees with the results of Stevens and Brookes.[9] This result implies that vision maintains a representation of surface orientation that is veridical with the surface being viewed when the surface displays only nonconflicting pictorial cues, at least over small connected regions of a surface.

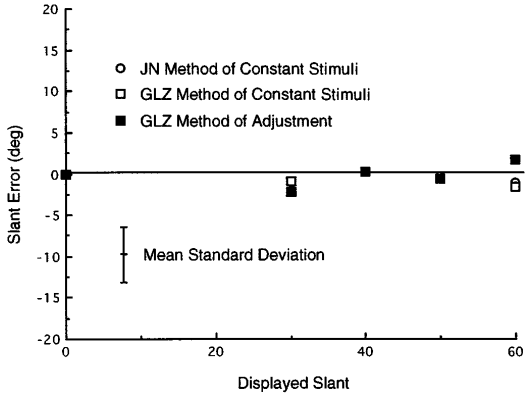

Fig. 8. Perceived slant error derived from Γ line-length matches, for three observers.

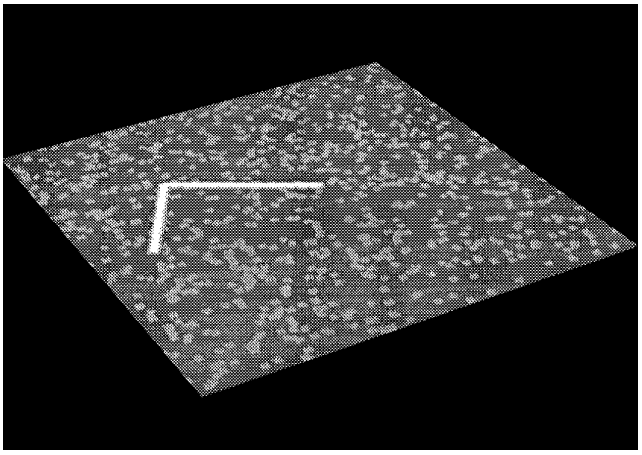

Fig. 9. Example of the stimulus used for experiment 3.

4. EXPERIMENT 3: ELIMINATING THE ASPECT-RATIO CUE

We considered two alternative strategies that subjects may have used to perform the line-length task that do not involve perception of surface slant. One method involves texture density; the other method involves aspect-ratio cue. In the texture-density strategy a subject could count the number of texture elements on the surface and adjust the probe such that both lines covered approximately the same number of texture elements. We evaluated the performance of the counting strategy in several simulated runs. The performance variability (associated with the random placement and density of the texture elements) was much greater than that of the human data. For this reason, and because the human subjects did not report finding a counting strategy useful, we conclude that the human subjects did not base their performance on this strategy.

The other strategy uses the aspect ratio of the surface boundary. Because the planar surface is a square and the test probe also forms one corner of a square, subjects might simply do the task by estimating the aspect ratio of the bounding surface and adjust the test probe to have the same aspect ratio.

It seemed unlikely that subjects were using this cue, given our findings in experiment 2 that subjects matched horizontal and vertical lines of unequal length in a plane of zero slant (horizontal–vertical illusion.) But it is possible that an aspect-ratio comparison could be based on a perceptual representation that is already distorted by the horizontal–vertical illusion.

We performed a third experiment, using a square plane oriented 30 deg with respect to the angle of tilt, as shown in Fig. 9. All the cues to slant were present in the image, but now they affected the sides of the plane and the texture differently than they affected the test probe. The foreshortening of the sides of the square plane was not the same as the foreshortening of the arm of the test probe, because the sides and the test probe were no longer parallel.

A. Method

This experiment was exactly like experiment 2; the method of adjustment was used. All the details are the same, except that the planar surface was rotated by 30 deg with respect to the tilt angle. A representative stimulus is shown in Fig. 9 for a slant angle of 60 deg.

B. Results and Discussion

The results are shown in Fig. 10. They show that the perceived slant calculated from the matched line lengths was accurate to within 2 deg of the veridical slant, with a standard deviation of 3.15 deg. The subjects demonstrated no evidence of recession to the frontal plane. The results clearly demonstrate that the subjects were not using the aspect ratio of the surface boundary to perform the line-length judgments.

We expected this result, because the horizontal–vertical illusion was present in all the data. Even though the sides of the plane were perfectly square, subjects seemed to ignore this and continue to set the vertical arm slightly shorter than the horizontal arm when the plane was frontal. If they were using the boundaries to make their length decisions, we would expect no illusion.

5. EXPERIMENT 4: REDUCTION OF THE AVAILABLE PICTORIAL CUES

In the previous experiments we demonstrated that subjects were capable of extracting surface slant from images that had texture and boundary cues. What happens to the perceived slant when one of the cues to surface slant is removed? In experiment 4 we removed either boundary-cue information or texture-cue information.

A. Method

The experiment was the same as experiment 2; the method of constant stimuli was used. Subjects were presented with a single plane with the Γ test probe painted on its surface. The slant of the surface was depicted either by boundary cues alone (the shape of the planar surface with respect to the background) or by texture cues alone (the shape of the elements and the density gradients). In the boundary-cue condition the plane was a continuous dark gray. Therefore the only cues to the plane's 3-D orientation were the perspective distortion and the aspect ratio of the boundary. In the texture-cue condition the same texture was viewed through a square aperture (6 cm × 6 cm). The surface of the plane was extended such that the boundary of the plane was always hidden by the aperture. Texture-element shape and the

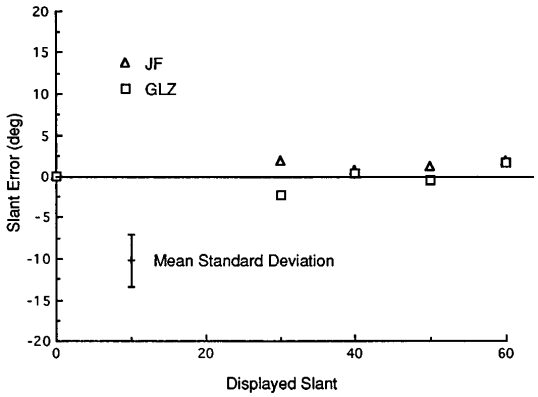

Fig. 10. Results of experiment 3 with use of the method of adjustment, for two observers.

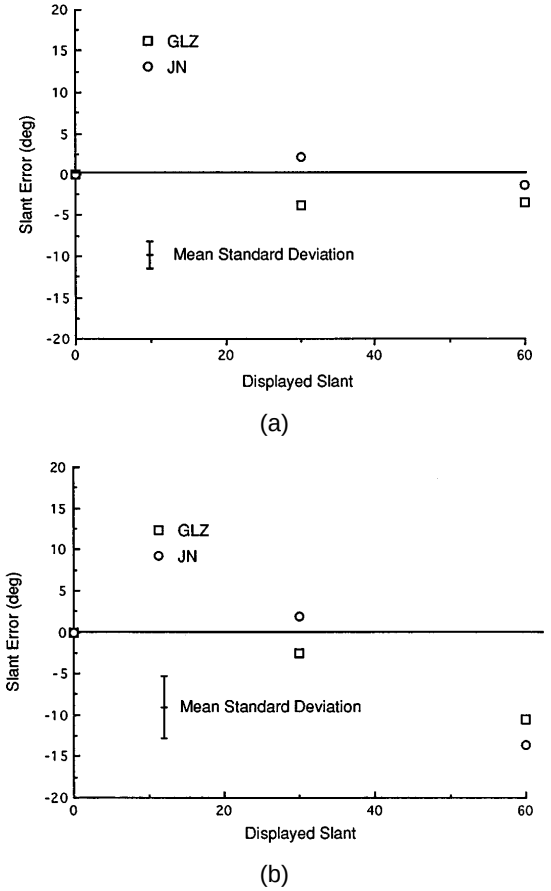

Fig. 11. Perceived slant angle for the conditions of (a) boundary cues alone and (b) texture cues alone, for two observers.

gradient of the texture were the available cues to spatial structure in this condition.

B. Results and Discussion

The results from this experiment are shown for two subjects in Fig. 11. With only the boundary cues available, subjects scored slightly worse, within 5 deg of the veridical slant. When the boundary-depth cues are presented alone and viewed monocularly, the surface slant is well defined, with few if any conflicting cues (e.g., the display frame) to suggest an alternative slant. Perception is sometimes called on for making judgments of surface slant in the absence of texture (e.g., for silhouettes).

For the condition with texture cues alone, subjects showed a dramatic underestimate of the slant angle at 60 deg. The underestimate of slant was 12–15 deg, with a doubling of the mean standard deviation to 3.7 deg. Subjects demonstrated dramatic recession to the frontal plane. This could mean that subjects are poor at computing the slant of a surface from pure texture cues alone. An alternative explanation is that the removal of a cue to the surface slant is perceived as a cue for the frontal position of the screen. The square aperture strongly suggested the frontal nature of the pattern being viewed. The underestimate of slant was much less dramatic at 30 deg. At 30 deg the difference between a square aperture and the shape of the boundary is small, and therefore the conflicting cue created by the aperture may not be so strong.

6. EXPERIMENT 5: PERFORMANCE UNDER EQUILUMINANT CONDITIONS

In experiment 4 we demonstrated that we could change performance on the line-length-matching task by modifying the pictorial cues. It has been demonstrated by many researchers that the perception of space from binocular disparity or motion parallax is degraded under equiluminant conditions.[10–12] If surface orientation derived from binocular disparity, motion parallax, and pictorial cues is maintained in a single representation, then we might expect orientation derived from pictorial cues to demonstrate a similar degradation at equiluminance. How do equiluminant conditions affect our perception of space from pictorial cues? In experiment 5 we examined this question by using only equiluminant pictorial cues to surface slant.

A. Method

There were two conditions in this experiment: a color condition and a luminance condition. The stimulus in both conditions was equally blurred so that the effects of chromatic aberration in the color condition were reduced. The blur was introduced by placement of a piece of frosted acetate as a filter that was flush with the screen. The extent of the blur was evaluated psychophysically by measurement of contrast thresholds for identifying grating orientation for 1 and 2 cycles/deg with the blurring filter. The filter attenuated the contrast sensitivity for 2 cycles/deg grating by a factor of 7 with respect to the contrast sensitivity measured for the 1-cycle/deg grating. We considered that the filter was effective at reducing components above 2 cycles/deg, which are the main source of luminance artifacts resulting from chromatic aberration.[13]

Subjects viewed the screen monocularly from a distance of 26 cm. When the plane was frontal it covered a visual angle of 17 deg. The texture elements covered ~1 deg of visual angle and were randomly placed across the surface of the plane. The texture elements and the plane were large so that they could be easily resolved at equiluminance.

In the luminance condition the lines were light gray, the texture elements were gray, the plane was dark gray, and the background was black, as in experiments 1–4. In the color condition the texture was green and the background was red. The plane itself was a yellowish color created by a mixture of green and red. The red of the background was kept at a constant level. The green and the yellow of the plane were found in the following way: the green luminance value used for 0% red–green (R–G) luminance contrast corresponded to the heterochromatic flicker photometry null point. The yellow of the plane was the point corresponding to halfway between the red of the background and the green of the texture. At 0% R–G luminance contrast the plane, the texture, and the background were all equiluminant.

The probe was a green–blue and was not equiluminant with the plane. It was visible at all R–G luminance contrasts. The R–G luminance contrast was varied around the zero point so that the static equiluminant point of the individual subject was sure to be included. Negative values of R–G luminance contrast correspond to a stimu-

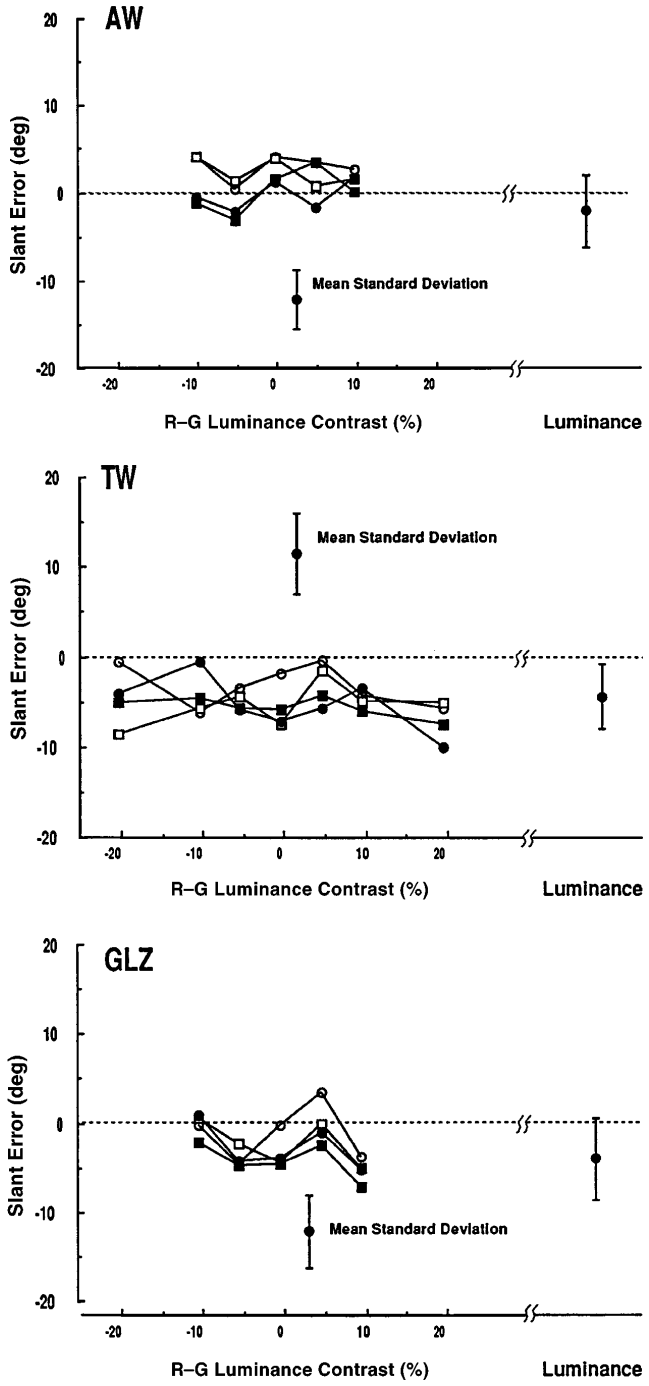

Fig. 12. Residual error in perceived slant angle in the color condition, for three observers. Slant residuals: open circles, 30 deg; open squares, 40 deg; filled circles, 50 deg; filled squares, 60 deg.

lus with more green than red. The task was the same as in experiment 2: matching the length of the variable and the standard lines of the Γ probe. Data were gathered by the method of adjustment.

B. Results and Discussion

Figure 12 shows data for all three subjects in both the luminance and the color conditions. At the far right of each graph are the average results from the blurred luminance condition. All three subjects showed a slight recession to the frontal plane. All three subjects were within 5 deg of the veridical slant. The slightly poorer performance on this task when it is compared with the results from experiment 2 is likely due to the addition of a blur screen, the increase in size of the texture elements, and the reduction in density of the texture elements. These results are in agreement with the results found in experiment 4 in the boundary-cues-only condition.

The results from the color condition are shown by the symbols in the left half of each graph. The horizontal axis is the R–G luminance contrast. Each symbol displayed on this graph corresponds to a measurement at a different slant angle (see caption). There are two things to note from the graph. First, the data points at each angle are tightly clustered; they show no characteristic changes with variation in luminance contrast. The performance at equiluminance was as good as it was with high luminance contrast. This is unlike the case for binocular disparity and motion parallax, in which performance improves rapidly as luminance contrast departs from equiluminance.[10–12] Second, two of the subjects show little or no recession to the frontal plane, and one subject showed some recession to the frontal plane. One possible explanation for the recession is that the blurring of the boundary and the contour of the texture created a conflicting cue, making the plane seem more frontal than it actually was.

If equiluminant conditions had a great effect on performance, we would expect to find a marked change at some point along these curves. We see that luminance contrast is not necessary for good slant perception from pictorial cues; color contrast alone can support good perception of surface orientation.

Our results show that color alone could mediate the perception of slant at equiluminance from boundary cues and texture cues. There was not a great loss in performance at equiluminance. This is in sharp contrast to the case for binocular disparity and motion parallax, in which performance degrades rapidly as luminance contrast approaches equiluminance. The one caveat is that the pictorial cues be large enough for the color pathway to resolve.

7. CONCLUSION

Unlike in most previous studies of slant perception, we found that pictorial cues can support an accurate perception of the spatial orientation of a planar surface. When the experimental probe was embedded in the surface of the plane under observation, subjects showed no perceived recession to the frontal plane. We did find a recession to the frontal plane in our first experiment, when the measurement required the comparison of two separate planes. The combination of these two findings suggests that accurate slant perception is limited to local surface area and is not necessarily maintained in register with the global map of depth. This is consistent with the findings of Stevens and Brookes[9] that subjects were good at estimating the slant of a surface from contour lines. In their experiments, accuracy with respect to pictorial cues was not maintained across the whole undulating figure when binocular disparity was used as an absolute measure of depth. We conclude from our findings and those of others

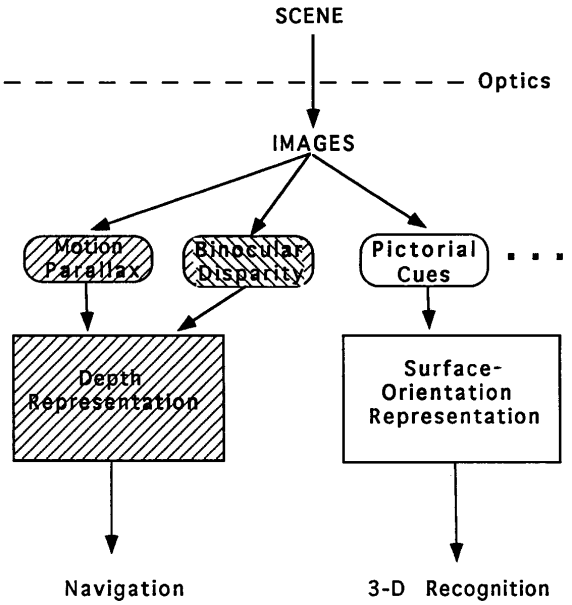

Fig. 13. Block diagram of proposed model of perceived representation of space.

that an observer's knowledge of the absolute position of the plane is quite loose for disjoint surfaces. This alone could account for many of the discrepancies found in the literature on slant perception.

Accurate measurement of perceived surface slant through surface comparison may be impossible to achieve. For a subject to derive perceived slant by comparing two planes implies knowledge of the perceived distance from the subject to each of the planes. Accurate knowledge of this kind is probably not available.[8] To fix the perceived distances with use of another cue such as binocular disparity implies that the representation of space is unitary and is maintained in register for all depth cues and that information from one depth cue can somehow drive the representation to a common value. This is probably not possible, either. Dichoptic measurement of surface slant compared with surface slant defined by pictorial cues will run into the same problem. If the representation of slant from one cue is held independently from another, cross-cue measurements will not necessarily accurately reflect the perceived slant.

Another aspect of these results reflects directly on experiments with conflicting depth cues. Many of the previous studies[14–16] implicitly assumed that there is one representation of depth derived from all visual cues. Often they used one depth cue (usually stereo disparity) as an anchor point in this representation of depth. This is not unreasonable. After all, if we have access to the interocular distance and the vergence angle, disparity is a direct measure of absolute depth. If, on the other hand, we are armed with two representations of depth derived from visual cues, these representations may be only loosely coupled or possible completely independent. Teasing them apart under normal visual experience would be difficult, because the depth cues themselves are often closely interrelated in the image. If we look at these experiments in this light, many confounding results can be reconciled.

We are left with a perplexing problem: How do we perceive space and organize that perception into a representation of the world? We propose the following speculative model, which may help to reconcile some of these contradictory findings on spatial representation. In this model the visual system maintains at least two functional representations of space: the depth representation and the surface-orientation representation. Each is derived from specific cues to the spatial structure of the scene and is used to perform specific visual tasks. A block diagram of the model is shown in Fig. 13.

The surface-orientation representation is derived primarily from pictorial cues. It is consistent with respect to these cues (i.e., no recession to the frontal plane) and very local (small connected surface patches). The representation is used for 3-D object recognition or for interpreting surface markings.

The depth representation is derived primarily from motion parallax cues, with some help from binocular disparity. It is global and very malleable (it will change dynamically). This representation is used for navigation through space. The main function of this representation is to tie together the glimpses of the scene gathered from each saccade into a single global perception of the world and to reference the observer spatially within that world. There is some recent evidence that the representations of motion parallax and stereo disparity are tied closely together.[17]

For the most part the two representations are in register with each other (i.e., surface orientation is equal to the derivative of the depth representation) simply because the orientation and the depth of the surfaces in the scene are in register. The independence of these representations becomes apparent in situations in which the pictorial and disparity cues are in direct conflict. One example of this might be in a still photograph. We are simultaneously able to see and understand the orientation of objects with respect to one another within the frame of the picture and to perceive the flat surface of the photograph itself.

Since we are constantly in motion and our eyes are able to capture only a small fraction of the available information presented in the image at a single instant, it is reasonable that motion parallax and binocular disparity play the role of tying the pieces of space together. Each of those pieces does not need to maintain close registration between the orientation of the patches of surface and the depth in the surrounding world, because we are always moving in that world. Pictorial cues, on the other hand, are involved in relaying information about the orientation of the local surface patches. The representation of space extracted from pictorial cues alone may be critical for visual tasks such as recognition and reading.

ACKNOWLEDGMENTS

This research was supported by a grant from the Center for Research in Learning Perception and Cognition at the University of Minnesota and National Science Foundation grant NSF/LEQSF(1992–96)-ADP-04 to G. L. Zimmerman, grant EY02857 from the National Eye Institute to Gordon Legge, and a Harvard Seed Grant to Patrick Cavanagh.

REFERENCES

1. J. M. Foley, "Binocular distance perception," Psychol. Rev. **87,** 411–434 (1980).
2. J. J. Gibson, *The Perception of the Visual World* (Houghton Mifflin, Boston, Mass., 1950).
3. H. R. Flock, "Three theoretical views of slant perception," Psychol. Bull. **62,** 110–121 (1964).
4. A. L. Kraft and W. A. Winnick, "The effect of pattern and texture gradients on slant and shape judgments," Percept. Psychophys. **2,** 141–147 (1967).
5. R. R. Rosinski, D. Degelman, and T. Mulholland, "Intermodal relationships in children's perception," Child Devel. **49,** 1089–1095 (1978).
6. J. A. Perrone, "Visual slant underestimation: a general model," Perception **11,** 641–654 (1982).
7. L. Matin and W. Li, "A single pitched line in darkness controls the elevation of visually perceived eye level," Invest. Ophthalmol. Vis. Sci. Suppl. **30,** 506 (1989).
8. J. T. Todd and F. D. Reichel, "Ordinal structure in the visual perception and cognition of smoothly curved surfaces," Psychol. Rev. **96,** 643–657 (1989).
9. K. A. Stevens and A. Brookes, "Probing depth in monocular images," Biol. Cybern. **56,** 355–366 (1987).
10. P. Cavanagh, C. W. Tyler, and O. E. Favreau, "Perceived velocity of moving chromatic gratings," J. Opt. Soc. Am. A **1,** 893–899 (1984).
11. J. Rivest, P. Cavanagh, and S. Saida, "Perception of depth from motion parallax at equiluminance," Invest. Ophthalmol. Vis. Sci. Suppl. **31,** 89 (1990).
12. C. M. de Weert, "Colour contours and stereopsis," Vision Res. **19,** 555–564 (1979).
13. P. Cavanagh and S. Anstis, "The contributions of colour to motion in normal and color-deficient observers," Vision Res. **31,** 2109–2148 (1991).
14. H. H. Bultoff and H. A. Mallot, "Integration of depth modules: stereo and shading," J. Opt. Soc. Am. A **5,** 1749–1758 (1988).
15. D. Buckley and J. P. Frisby, "Interaction of stereo, texture, and outline cues in the shape perception of three-dimensional ridges," Vision Res. **33,** 919–933 (1993).
16. M. J. Young, M. S. Landy, and L. T. Maloney, "A perturbation analysis of depth perception from combinations of texture and motion cues," Vision Res. **33,** 2685–2696 (1993).
17. M Nawrot and R. Blake, "On the perceptual identity of dynamic stereopsis and kinetic depth," Vision Res. **33,** 1561–1571 (1993).