Three Gradients and the Perception of Flat and Curved Surfaces

James E. Cutting and Robert T. Millard
Cornell University

SUMMARY

Researchers of visual perception have long been interested in surfaces. Most psychologists have been interested in the perceived slant of a surface and in the gradients that purportedly specify it. *Slant* is the angle between the line of sight and the tangent to the planar surface at any point, also called the *surface normal. Gradients* are the sources of information that grade, or change, with visual angle as one looks from one's feet upward to the horizon. The present article explores three gradients—perspective, compression, and density— and the phenomenal impression of flat and curved surfaces. The *perspective gradient* is measured at right angles to the axis of tilt at any point in the optic array; that is, when looking down a hallway at the tiles of a floor receding in the distance, perspective is measured by the x-axis width of each tile projected on the image plane orthogonal to the line of sight. The *compression gradient* is the ratio of y/x axis measures on the projected plane. The *density gradient* is measured by the number of tiles per unit solid visual angle. For flat surfaces and many others, perspective and compression gradients decrease with distance, and the density gradient increases. We discuss the manner in which these gradients change for various types of surfaces. Each gradient is founded on a different assumption about textures on the surfaces around us.

In Experiment 1, viewers assessed the three-dimensional character of projections of flat and curved surfaces receding in the distance. They made pairwise judgments of preference and of dissimilarity among eight stimuli in each of four sets. The presence of each gradient was manipulated orthogonally such that each stimulus had zero, one, two, or three gradients appropriate for either a flat surface or a curved surface. Judgments were made for surfaces with both regularly shaped and irregularly shaped textures scattered on them. All viewer assessments were then scaled in one dimension. Multiple correlation and regression on the scale values revealed that greater than 98% of the variance in scale values was accounted for by the gradients. For the flat surfaces a mean of 65% of the variance was accounted for by the perspective gradient, 28% by the density gradient, and 6% by the compression gradient. For curved surfaces, on the other hand, a mean of 96% of the variance was accounted for by the compression gradient, and less than 2% by either the perspective gradient or the density gradient. There were no differences between results for surfaces with regularly shaped and irregularly shaped textures, demonstrating remarkable tolerance of the visual system for statistical variation.

The differential results for the flat and curved surfaces suggest independent channels of information that are available in the optic array to observers for their use at different times and in different situations. We argue that perspective information seems to be most important for flatness judgments because that information is a component of an invariant available to viewers about flat surfaces. We also argue that compression is important for curvature judgments because it reveals potential nonmonotonicities in change of slant, the angle between line of sight and a line orthogonal to the local surface plane. In Experiment 2 we show that when the height of a curved surface is diminished enough to create a nearly monotonic compression function, viewers cannot distinguish such a surface from one that is flat.

Finally, we suggest two things with regard to the existing literature on surface perception. First, although psychologists have been very interested in the perception of slant for a flat surface, we argue that slant is a largely irrelevant variable for the perception of flatness. That is, slant information is functionally related to compression gradients, and compression

gradients appear not to be used by perceivers in their judgments of flatness in surfaces. Instead, relative slants are important in the perception of curved surfaces, where compression accounts for almost all of the variance in viewers' judgments of curvature. Second, although researchers in the field of artificial intelligence have largely shunned the study of gradients as irrelevant to the perception of surfaces, we believe that our results suggest that the information in gradients is crucial for their perception.

How do we perceive the layout of surfaces? What information is available to us, and what information do we use? Such questions have intrigued psychologists for a third of a century (see Gibson, 1950) and they are increasingly central to artificial intelligence approaches to vision (see Bajcsy & Lieberman, 1976; Marr, 1982; McArthur, 1982; Stevens, 1981a; Witkin, 1981). Gibson and his followers took the position that we perceive surfaces with regard to *slant*—the angle of incidence between the line of sight and a line orthogonal to any particular point on the surface of regard, called the *surface normal*. It is slant, according to Gibson (1950; but see 1979, p. 196), that gives rise to various gradients of texture and that promotes the perception of surfaces receding in depth.[1]

Continuing interest in perception of slant has generated a large literature (see, e.g., Attneave, 1972; Beck & Gibson, 1955; Braunstein & Payne, 1969; Clark, Smith, & Rabe, 1955; Epstein, Bontrager, & Park, 1962; Flock, Graves, Tenney, & Stephenson, 1967; Freeman, 1966; Gillam, 1970; Olson, 1974; Perrone, 1980; Phillips, 1970; and Rosinski & Levine, 1976). This literature generally shows two things: (a) observers are not very good at judging absolute slant, and (b) they are quite good at judging differences in slant. The first of these findings seems unsurprising and not particularly relevant to the perception of surfaces; the second is relevant, but has not, in our opinion, been applied to surface perception in an appropriate way.

With regard to the first point, for example, we believe that observers do not often need information about absolute slant. When looking out onto a flat terrain, one finds that the slant of the surface of support is perpendicular to the line of sight at one's feet, and that it grades uniformly to the horizon where it becomes parallel to the line of sight. At all points, the surface is flat, but slant is changing. It is difficult to understand how computation (or pick up) of absolute slant would be relevant to the perception of the shape of the surface without other information, such as where particular portions of surface are with respect to the observer. Moreover, the observer needs to perceive only that relative slants and tilts at different points in the visual field give rise to the shape of that surface and to the relative distances of various points on it.

With regard to the second point, the difference between slants of two patches of surface is also insufficient evidence in itself to determine whether these two parts of a surface are coplanar or noncoplanar. Information about spatial arrangement must also be available. In other words, the same change in slant, depending on overall spatial arrangement and on viewpoint, could indicate a flat surface, a curved surface, or even two unrelated surfaces. Slant, as it turns out, appears to be crucial only to the perception of certain kinds of surfaces, and we will return to its discussion later. In essence, we demonstrate that slant is largely irrelevant to the perception of flatness in surfaces.

A second and related literature has not measured perceptual sensitivity to slant. Instead, it has dealt with the phenomenal impression of three-dimensionality of receding surfaces (see, e.g., Attneave & Olson, 1966; Braunstein, 1976; Marr, 1982; Stevens, 1981a;

[1] Researchers now realize that *tilt*—the angle between the surface normal and true vertical as measured on a projection plane orthogonal to the line of sight—is a separate and equally important variable (Marr, 1982; Stevens, 1981a, 1983). Nevertheless, it is slant that has captured the attention of psychologists.

Vickers, 1971; Witkin, 1981). It is this literature, we believe, that is a more appropriate starting point in trying to answer the questions posed at the beginning of this article—how we perceive surfaces, what information is available, and what information we use. Typically in these studies, researchers employ systematic variation in surface textures to study the effect on perception of surfaces receding in depth. This variation grades with visual angle (or height in field where 90° is at the horizon and 0° is at one's feet), and gives rise to the study of gradients.

Gibson (1950, p. 78) was the first to discuss the perception of surfaces in terms of "stimulus gradients of the density of texture and the size of objects." This long and cumbersome phrase was then shortened to "density gradients" or sometimes "texture gradients." These gradients were taken, by Gibson, to be the primary source of information for the perception of surfaces. But gradients have speciated since Gibson first wrote of them. We now know, for example, that the density gradient, per se, is not very effective in revealing a surface receding in depth (Braunstein, 1976; Marr, 1982; Stevens, 1981a), but this jumps us too far ahead in our discussion.

Assumptions and Gradients

First, we must consider the nature of the textures that pepper the surfaces around us. Whether we are talking about walls, floors, greenswards, parking lots, tundras, deserts, or ocean surfaces, three properties that generally hold are the following:

Assumption 1: Texture elements are roughly the same absolute size.

Assumption 2: Texture elements lie flat or nearly flat on the plane.

Assumption 3: Texture elements have roughly uniform, absolute spacing.

These are the properties that give rise to systematic change in the optic array. Study of flat surfaces has revealed many sources of information that grade with visual angle. These are all separate texture gradients. Six such gradients are most often discussed. Each is typically measured vertically in the picture plane under the assumptions that the surface has

neither tilt nor horizontal perturbations. The first is the *perspective gradient,* and it follows from Assumption 1. Perspective, sometimes called *scaling,* can be measured in several ways. The most secure is the horizontally measured relative optic angle for any given texture. This gradient changes as a direct function of three parameters: d or the distance of a particular texture from the observer's feet, e or the eye height of the observer, and t or the size (radius) of any given texture in the real world. Of course, the additional assumption that textures are roughly circular is also being made. The optic angle P for the perspective of any given texture along the line of slant is expressed as the following:

$$P = 2 \cdot \arctan[t/(d^2 + e^2)^{1/2}]. \quad (1)$$

Convergence, or linear perspective, is a special case of this gradient, an idea that we will return to in the discussion of perspective and flatness. The second is the *compression gradient,* which is sometimes called *foreshortening.* It follows from Assumption 2. Compression, the projectively vertical optic angle C for any given texture element along the line of slant, is a function of these same parameters:

$$C = \arctan[(d + t)/e]$$
$$- \arctan[(d - t)/e]. \quad (2)$$

It is customary to measure compression against perspective, hence the compression gradients used to generate stimuli in this study are actually C/P at a given distance and eye height. But because Equation 2 in its present form will be helpful throughout this presentation, we will use it to stand for the independent information in compression gradients. The third is the *density gradient,* which results from Assumption 3. Density is the number of texture elements per unit solid visual angle. Ignoring the difference between spherical and planar geometry and assuming that the textures are equally spaced at an average of $2n$ units apart in all directions, measured from texture center to texture center, the relative density gradient D at any point along the line of sight can be specified as the following:

$$D = 1/\{\arctan[(d + n)/e]$$
$$- \arctan[(d - n)/e]\}$$
$$\times 2 \arctan[n/(d^2 + e^2)^{1/2}]. \quad (3)$$

Notice that if t and n are systematically related, D is a function of P and C. A fourth gradient is the *size gradient* but this is always a straightforward multiplication of P and C, a fifth is the *motion gradient* (see Braunstein, 1976; Gibson & Carel, 1952), and a sixth is the *intensity gradient* (see Horn, 1975; Marr, 1982; Todd & Mingolla, 1983). The latter two are beyond the scope of this article. The first three—compression, density, and perspective—can be specified independently for a static surface, and it is these gradients that concern us.

It is generally known that the perspective gradient is more potent than compression (e.g., Rosinski & Levine, 1976; Vickers, 1971) for the perception of a flat surface receding in the distance, and that the density gradient is neither necessary (Marr, 1982, p. 237) nor sufficient (Braunstein, 1976, p. 48). But we do not generally know the full ordinal rankings of these three gradients in their use for perception. One of the goals of this study is to determine the relative importance of each of these gradients in the judgment of flat surfaces—with the expected outcome that the perspective gradient would be more important than the other two. Moreover, importance is assessed not only in ordinal terms, but in proportional terms as well through the use of various scaling and regression techniques. A second goal of this study is to investigate the relative importance of these three gradients for the perception of curved surfaces. Little is known about the perception of these gradients in nonflat conditions (although Todd & Mingolla, 1983, have studied intensity gradients for curvature). One might expect that compression would be relatively more important for curved than for flat surfaces, but no other clues to possible outcomes seem available. What follows is an empirical analysis of the relative import of three gradients in the perception of flat and curved surfaces.

Experiment 1: Perceptual Assay of Flat and Curved Surfaces

Method

Ten viewers participated in four tasks: the 2 authors, and 8 members of the Cornell community who were naive to the notion of gradients and who were paid for their services. Two tasks entailed judgments of the flatness of textured surfaces receding in the distance, and two entailed

the judgments of the curvature of textured surfaces receding in the distance. Crossed with this flat/curved distinction was another: In two tasks the viewers were presented with surfaces peppered with regular octagons, and in the other two they saw surfaces with irregular octagons. The four tasks were called flat/regular, flat/irregular, curved/regular, and curved/irregular. Tasks were self-paced and took about 35 min each to complete. The order of participation was balanced in a partial Latin Square design (with 2 observers in each of two orders, and 3 in each of two others). Viewers participated individually for two sessions, each lasting 1 hr 15 min. Each task was preceded by a four-item practice sequence to familiarize viewers with the stimulus range and with the responses they were to make.

Each task consisted of 56 trials: all possible pairwise comparisons among eight different stimuli (28) taken in both left–right (L–R) and right–left (R–L) configurations. Each viewer was presented with a different random order of trials. On a given trial, the viewer made two responses of stimulus comparison. First, he or she indicated which member of the pair (L or R) looked more like a flat (or curved) surface receding in the distance. This judgment was called the *preference judgment,* and formed the basis of Case 5 Thurstonian scaling, discussed later. Second, he or she indicated the degree (1 to 9) to which the preferred stimulus revealed flatness (or curvature) over the nonpreferred stimulus. This judgment was called a *dissimilarity judgment,* and formed the basis of metric multidimensional scaling, discussed later. Viewers typed their responses on a console, and individual data files were arranged and stored by computer. These responses, which were clearly inferential and not direct on the part of the observer, were used to make interpretations about phenomenal impressions of the stimuli. Thus, direct perception (Gibson, 1979) was not being tested; only the information that could serve direct perception was being assessed.

Stimuli were generated on a Hewlett–Packard (HP) 1000L computer, and displayed in an HP1350S vector-plotting display system. Each stimulus pair was generated uniquely for each trial, for each condition, and for each subject. They were shown side by side to the viewer, with a vertical line between them. Since the experimental room was moderately lit, the sides of the monitor could be seen as well, forming a frame around the screen image. Thus, the stimuli looked like representations of depth rather than real depth. Sample stimuli for each of the eight trial types in each of the four conditions are shown in Figures 1 and 2. The eight stimuli for the flat/regular and flat/irregular tasks are shown in Figure 1, and those for the curved/regular and curved/irregular tasks are shown in Figure 2. In each case, Stimulus 1 indicated a stimulus with no gradients appropriate to the task judgment. Stimuli 2–4 have one appropriate gradient apiece: Stimulus 2 shows only a compression gradient, Stimulus 3 shows only a density gradient, and Stimulus 4 shows only a perspective gradient. Stimuli 5–7 have two appropriate gradients: Stimulus 5 has both compression and density, Stimulus 6 has both compression and perspective, and Stimulus 7 has both density and perspective. Stimulus 8 has all three appropriate gradients. In Figure 1, those stimuli without a particular gradient appropriate for flatness have a distribution of textures that is uniform from the bottom to the top of the panel. Technically, this means that the gradient is appropriate for a surface orthogonal to the line of sight
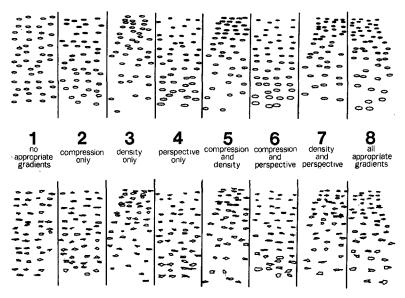
Figure 1. At the top are eight sample stimuli representing flat surfaces peppered with regular textures, and at the bottom are eight with irregular textures. (Stimuli 1 have no gradients appropriate for a receding flat surface, instead they are those of a surface orthogonal to the line of sight; Stimuli 2–4 have one gradient apiece, with Stimuli 2 having a compression gradient, Stimuli 3 a density gradient, and Stimuli 4 a perspective gradient; Stimuli 5–7 have two appropriate gradients, with Stimuli 5 having both compression and density, Stimuli 6 compression and perspective, and Stimuli 7 density and perspective; and Stimuli 8 have all three gradients. These representations, because of their reduction in size, are not nearly so impressive as surfaces receding in depth as the original stimuli. Because of the use of polar projection and because of the reduction in size, the computed station point for each stimulus would be about 2 in. [5 cm] away from this page.)
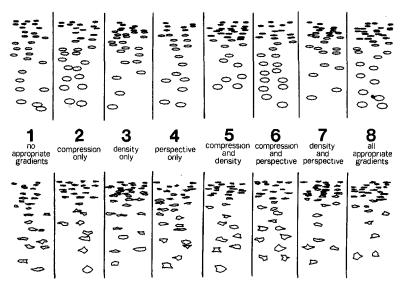

Figure 2. The same arrangements as in Figure 1, but for curved surfaces. (Those at the top have regular textures and those at the bottom irregular textures. Inappropriate gradients are those for a flat, receding surface. Thus, Stimuli 1 in this figure are the same as Stimuli 8 in Figure 1. Again, these sample stimuli are much reduced and hence less impressive than those seen by the participants in the experiment.)

(such as a slightly curved wall) and inappropriate for a horizontal surface receding from view. Since the regular octagons are slanted, the textures in Stimulus 1 may be those on a venetian blind. Again, one should note that compression values are considered as C/P, not simply C in Equation 2. In Figure 2, those stimuli without a particular gradient appropriate for curvature have a gradient that is appropriate for the flat surface of Figure 1. Thus, in Figure 1 the conflict of information in Stimuli 2–7 is between gradients that specify horizontal flatness versus near vertical flatness, whereas in Figure 2 the conflict is between gradients appropriate for curvature and those appropriate for horizontal flatness. In other words, Stimuli 1 in Figure 2 are the same (in terms of gradients) as Stimuli 8 in Figure 1. Because of their reduction in size, the sample stimuli of Figures 1 and 2 are not nearly as convincing as those viewed by the participants.
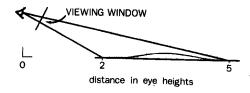
The projective geometry of the viewing situation was as follows: The viewer sat so that he or she saw the stimuli from the projected station point. If one considers the unit of distance from the viewer's eye to the surface of the projected plane as 1 eye height, then the front of that plane was at a distance of 2.0 eye heights from the observer's feet and the back of that plane at 5.0 eye heights. The distance across the front of the surface seen through the visible window was 0.16 eye heights, and the distance across the back of the seen surface was 0.36 eye heights. For flat surfaces all textures were viewed from a uniform eye height; for curved surfaces, the curvature was convex, following an inverse cosinusoid ($1 - \cos \beta$) with an excursion depth of 0.14 eye heights. That is, the front edge and back edges of the viewed surface were at 1.0 eye heights, but in between the surface sloped upward such that the middle was 0.86 relative eye heights below the station point. For an observer 1.83 m (6 ft) tall, with an eye height of about 1.7 m, the height of the cosinusoidal hill was 0.238 m (or 9.37 in.). This hill was one dimensional, with variation only from front to back (or bottom to top in the picture plane). There was no tilt and no variation in tilt. The relations are shown in the top panel of Figure 3. Notice that there were no occlusions.
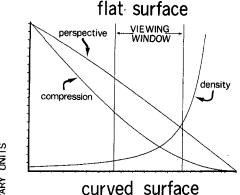
To avoid the powerful cues of linear perspective, textures (regular or irregular octagons) were placed on the surfaces at random locations within cells on a grid. Each cell on the surface was a rectangle. These cells projected as trapezoids on the scope face for flat surfaces and curved surfaces alike. For flat surfaces, regular octagons (projected as nearly ellipsoidal) were 0.015 eye heights in diameter; irregular octagons were formed by generating a random number between 0.01 and 0.02 eye heights and using that value as a radius (r value) for each of eight angles (θ values) spaced 45° around a circle. Thus, the irregular shapes had the mean radial statistics of the regular shapes, but no invariant properties per se. For the curved surfaces the texture-size statistics were twice as great, and their number about half as many. The rationale for this change was computational: Because each stimulus was computed uniquely for each trial and because the computations for a curved surface are much more complex, the intertrial interval would have been inordinately long—on the order of a minute—had the textures been as plentiful in the curved surface conditions.

If one considers the horizon to be at an optic angle of 90° and a point directly beneath one's eyes to be at 0°,

then the window through which the textured surfaces were seen was between 63° and 78°. Using textures at 45° (or the horizontal distance outward of 1 eye height) as an
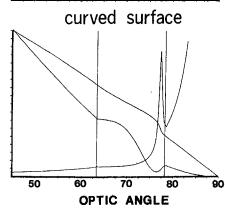

Figure 3. Geometry of the viewing situation and gradients for two surfaces. (The upper panel shows the viewed surface between 2 and 5 eye heights away from the observer, subtending the optic array between 63° and 79°, if the horizon is considered to be at 90°. The hill for the curved surface is 0.14 eye heights in elevation. The middle panel shows the relative values of each of the three gradients through the optic array from 45°, 1 eye height distant, to 90° for flat surfaces. In the text this is called optic angle α. Vertical markers indicate the region within the window of the display scope. All gradients are arbitrarily normalized at 45°. The bottom panel shows the relative values of the three gradients for the curved surfaces. What is labeled as compression is the function generated by Equation 2. True compression is actually C/P, or the outcome of Equation 2 divided by that of Equation 1.)

arbitrary anchor for calculating gradient values, the change in the C, D, and P gradients for a flat surface can be seen in the middle panel of Figure 3. These are generated from Equations 2, 3, and 1, respectively, and assume no variation in size, flatness, or density of textures.[2] Allowing for random normal variation in these measures, which is found for density in all four tasks and found for compression and perspective for the two tasks involving irregular textures, we might have introduced confidence bands around these functions. Thus, one could consider these gradient functions as broad smears with only their central tendencies shown in Figure 3. Notice that all three are different, but within the window used in these experiments the change in each of the gradients is roughly comparable. The gradients for the curved surface are shown in the bottom panel of Figure 3. Here one can see that the change in the gradients is rather marked for compression and density, but less so for perspective. On the basis of these plots, one might expect that all three gradients would be about equally important for judging the flatness of a surface within the range of 63° and 78°, and that C and D would be relatively more important than P in judging the curvature of surfaces. The results of the following experiment show that neither of these predictions is obtained.

Analysis

The data for each subject in each task were assembled into two 8 × 8 matrices with major diagonals missing: one matrix for the preference data and one for the dissimilarity data. Consider first the preference data. Phi coefficients were computed for the 56 judgments for each of the 10 different subjects for each task. Mean phi coefficients were .593 for the flat/regular task, .468 for the flat/irregular task, .336 for the curved/regular task, and .319 for the curved/irregular task. Since these means are all reliable ($p < .025$ or better), one can see that there is good correspondence among the viewers in their preference judgments on all tasks.

The 10 individual preference matrices were folded and then averaged to create a single half-matrix for each of the four tasks. Those matrices were then subjected to Case 5 Thurstonian scaling (see, e.g., Baird & Noma, 1978) in order to obtain one-dimensional preference scale values for the eight different types of stimuli in each task. The average discrepancies for each scaling solution were quite low, indicating (a) an excellent fit of the data by chi-square goodness of fit tests and (b) additivity of all judgments. The average discrepancy for the flat/regular task data was .07, $\chi^2(21) = 21.09$, $p > .35$; that for the flat/irregular task data was .05, $\chi^2(21) = 6.59$, $p > .99$; that for the curved/regular data .07, $\chi^2(21) = 9.84$, $p > .98$; and that for the curved/irregular data was .07, $\chi^2(21) = 8.79$, $p > .99$. The scale coordinates for the four tasks are given in Table 1.

The dissimilarity data were treated differently. The individual matrices were first folded across the diagonal, averaging those judgments of dissimilarity involving the same stimuli provided they passed a consistency check. This check can be described as follows: Consider a trial presenting Stimulus 2 on the left side and Stimulus 3 on the right as a 2/3 trial, and the reverse configuration as a 3/2 trial. If the observer responded L4 on the 2/3 trial and R7 on the 3/2 trial, then the observer consistently chose Stimulus 2 over Stimulus 3, and the dissimilarities would have been averaged and would have equaled 5.5. If, on the other hand, the observer responded L4 on the 2/3 trial and L7 on the 3/2 trial, then the observer was inconsistent and preferred Stimulus 2 once and Stimulus 3 once. In such cases the dissimilarity judgments would have been considered the absolute value of the difference, divided by 2: Thus $|4 - 7|/2 = |7 - 4|/2 = 1.5$. The data in the resulting half-matrices were then normalized to reduce the effect of individual differences in the use of the 1-to-9 rating scale.

To ensure that individual variation was minimal, the half-matrices for all subjects were intercorrelated for each task. Mean correlation for the flat/regular task dissimilarities was .842; that for the flat/irregular data was .781; that for the curved/regular data, .729; and that for the curved/irregular data, .692. Since these means are all reliable ($p < .001$ or better), one can see that the dissimilarity data also have good intersubject correspondence.

The 10 individual half-matrices were then averaged to create a single half-matrix for each of the four tasks. Those matrices were then subjected to multidimensional scaling, with

---

[2] Notice that what is indicated as compression in the lower two panels of the figure stems from Equation 2 for the flat surface and a modification of it for the curved surface. Compression, as defined for the stimuli used in these experiments, is divided by perspective. The two functions are separated here because, for reasons that are not analytically obvious to us, $C/P = kP$, or C (of Equation 2) = kP^2 (of Equation 1). In other words, the perspective (P) and compression (C/P) functions would be identical.

Table 1

Resulting Coordinates for Thurstonian Scaling of Preferences and Metric One-Dimensional Scaling of Dissimilarities

				Flat surface				Curved surface			
			Regular		Irregular		Regular		Irregular		
	Gradient											
	C	*D*	*P*	Prefer-ence	Dissimi-larity	Prefer-ence	Dissimi-larity	Prefer-ence	Dissimi-larity	Prefer-ence	Dissimi-larity
1.	0	0	0	2.902	1.361	2.379	0.824	1.347	1.268	1.360	0.992
2.	1	0	0	2.392	1.165	2.106	0.595	0.196	−0.883	0.087	−0.810
3.	0	1	0	1.941	0.626	1.592	0.327	0.974	0.913	1.178	1.156
4.	0	0	1	1.447	−0.149	1.225	0.091	1.225	1.005	1.406	1.058
5.	1	1	0	1.394	0.244	1.022	0.104	0.000	−0.767	0.287	−1.202
6.	1	0	1	0.895	−0.451	0.907	−0.034	−0.021	−1.002	0.189	−0.918
7.	0	1	1	0.754	−1.076	0.448	−0.674	1.064	0.742	1.052	0.751
8.	1	1	1	0.000	−1.719	0.000	−1.026	−0.108	−1.276	0.000	−1.028

Note. *C* stands for the compression gradient (actually *C/P*), *D* for the density gradient, and *P* for the perspective gradient. Presence of an appropriate gradient is indicated by 1, inappropriateness by 0. See the text and captions for Figures 1 and 2 for definitions of appropriateness. For both preference and dissimilarity values, the less the number, the more the stimulus revealed either a flat or curved surface.

the idea that solutions in one dimension would be optimal. Nonmetric scaling was attempted at first, but this led to several badly degenerate solutions. Metric scaling was then tried so that small dissimilarities would not be underemphasized. Both Equation 1 and Equation 2 stresses were employed. Stress Equation 1 led to somewhat more satisfactory one-dimensional solutions, where stress for the flat/regular task data was .20, that for the flat/irregular data was .33, that for the curved/regular data was .22, and that for the curved/irregular data was .19. Although direct comparison with the fits of the Thurstonian scaling is not possible, it is fairly clear that these data do not fit as well. With nonmetric scaling it is generally suggested (Kruskal, 1968) that with Stress Equation 1, one-dimensional stresses of .20 are poor. Because metric scaling will almost always lead to stress values that are quite a bit higher than corresponding nonmetric values, it is probably safe to say that these fits to the data are fair.[3] Equation 2 stresses for the four data sets given previously were .41, .47, .35, and .31, respectively. These values are roughly in keeping with the idea that Stress Equation 2 values are about twice those of Stress Equation 1. The scale coordinates for the one-dimensional Stress Equation 1 solutions for each task are given in Table 1.

Next, the stimulus coordinate values in the Thurstonian scaled solutions and in the one-dimensional metrically scaled solutions were used as dependent measures in multiple correlation and regression. Dummy values (1 = presence; 0 = absence) of the compression, density, and perspective gradients were used as independent variables, as shown in Table 1. Here, the presence or absence of each of the three gradients was used as a predictor of the scaled preference data and of the scaled dissimilarity data in each of the four tasks. Thus, eight different multiple correlations were computed, four for the preference data and four for the dissimilarity data. Each *R* was outstandingly high. The smallest of the eight *R* values was .992, meaning that more than 98% of the variance in each of the scaling solutions is accounted for by the three dichotomous variables: compression, density, and perspective. Thus, virtually none of the variance in the scale data is attributable to noise, and we can partial out the variance accounted for to the three independent variables.

---

[3] Shepard (1974, p. 385) suggested that too much emphasis has been placed on these evaluative labels. The utility of scaled solutions is not so much determined by the "badness of fit," as by interpretability. As we demonstrate, the scale solutions here are highly interpretable.

Since the three gradients occur in the eight stimuli in completely orthogonal manners (each correlates with each other, $r = .00$), the semipartial correlations are the same as the simple correlations. These values squared are the proportions of variance accounted for in the scaled solutions. They are shown in Table 2, and are the main results of Experiment 1. Because of the low discrepancies in the preference scale solutions and the moderate stress in the metric scale solutions, we felt justified in assuming independence of the three gradient measures.

Results and Discussion

There is a striking comparability among the four analyses for flat surfaces, and among the four analyses for curved surfaces. Regardless of whether one considers the preference data or the dissimilarity data, or whether one considers the data for the surfaces with regularly shaped textures or irregularly shaped textures, the results are virtually identical. For the judgments of flat surfaces, between half and three quarters of the variance is accounted for by the perspective gradient, between one quarter and one third by the density gradient, and very little by the compression gradient. For the curved surfaces, on the other hand, essentially all of the variance is accounted for by the compression gradient, and virtually none by either the density or perspective gradient.

The use of the metric one-dimensional scaling solutions, even though they had moderate stress, appears justified given the comparability of multiple correlation results with those for the preference data. The correlation of the variances accounted for in preference and dis-

similarity data is very high ($r = .992$). Given the relative independence of the two measures (the consistency check in the dissimilarity data keeps them from being completely independent), and given the comparability of results across regular and irregular textures ($r = .991$), the pattern of results becomes undeniable: The perspective gradient is by far the most important source of information for judging flat surfaces, but there is some contribution from the density gradient; and the compression gradient is the sole source of information used for judging curved surfaces. The negative correlation for the variances accounted for between flat and curved surfaces ($r = -.776$) suggests further that quite different information is being used in the two situations. Although it is difficult to separate weights from scale values in this type of design, the idea that the variances accounted for reflects perceptual impact is pursued.

There is some problem about the generality of the curved surface results. After all, every flat surface is equally flat, but curved surfaces can vary widely, and in two dimensions, not just the one dimension (projectively vertical) found in these stimuli. However, we think our results generalize to other (and perhaps all) curved surfaces given the modest size and smoothness of the slope used here, and given the power of the results. Once local orientation of the surface, or tilt, is determined (Marr, 1982; Stevens, 1981a, 1983), differences in compression along the axis of tilt will always be of use as the determinant of local curvature. Thus, compression should equally well be used to determine curves in walls and curves of a surface of support in both one and two dimensions (see, for example, Marr, 1982, p.

Table 2
Percentage of the Variance Accounted for by the Three Gradients in the Scale Solutions for Flat and Curved Surfaces

| | Flat surface | | | | | Curved surface | | | | |
| | Regular | | Irregular | | | Regular | | Irregular | | |
Gradient	Prefer-ence	Dissimi-larity	Prefer-ence	Dissimi-larity	M	Prefer-ence	Dissimi-larity	Prefer-ence	Dissimi-larity	M
Compression	11.4	3.6	7.4	3.0	6.4	95.4	96.4	95.1	97.8	96.1
Density	25.9	23.1	35.1	28.7	28.2	3.1	0.9	1.4	0.7	1.5
Perspective	62.3	72.8	56.6	66.5	64.6	0.6	1.7	0.4	0.1	0.7

217, Figure 3-56b). We investigate compression and curvature in more detail in a later discussion section.

Mismatch of gradient change and perception. The pattern of results for the flat and curved surfaces cannot be accounted for in terms of how much each of the gradients changes across the display, bottom to top. Consider first the flat surfaces. As shown in the middle panel of Figure 3, all of the gradients change a substantial amount, as generated by Equations 1–3. The ratio of initial-to-final values in the perspective gradients, going from the bottom to the top of the displays, is 2.27. The final-to-initial values of the density gradient are a ratio of 4.20, and the initial-to-final ratio for the compression gradient is also 2.27, when taken as C/P. Thus, whereas the perspective gradient changes no more than compression and changes less than density, it accounts for most of the variance in viewers' judgments.

This mismatch between results and the change in the gradients is also found for curved surfaces. Here, the comparison is between the value of each of the gradients for the curve at various optic angles compared with those values at the same angles for the flat surface. This entails a comparison between the lower and middle panels of Figure 3. At the inflection points, C is as much as 42% higher and 70% lower in value for the curved surfaces compared with the flat ones. But assuming textures are sampled throughout the density function (an assumption we will return to), density is as much as 206% higher and 39% lower in the same comparisons. Perspective is the gradient that changes least, and only in one direction: It is as much as 16% higher in the upper middle of the display. Most interesting is the fact that the compression gradient changes somewhat less than density, yet it accounts for essentially all of the variance in the viewers' responses.

The questions arise, of course, as to why the perspective gradient (P) should be the primary source of information for flatness and the compression gradient (C/P), for curvature. These questions are dealt with later. But first we emphasize that independent sources of information—the various gradients—are being used for different purposes in different situations. We believe that the real world provides a plenum of information that can be used in different ways and at different times. This study is a testimony to the fact that mere specification of the information available to a perceiver is a foolhardy course of action if we are to account for why we perceive things as we do. The study of invariants and of variants (like the gradients here) in perception is incomplete without the demonstration that these sources of information are used by the perceiver. Their differential use, as shown in this study, suggests that we must go beyond the information given and study how information is used (Hochberg, 1982).

Against strong inference. Finally, this kind of experimental analysis shows that we should be wary of certain demonstrational analyses like those, for example, done by Stevens (1981a; Marr, 1982; see also Braunstein, 1976). Stevens demonstrated that density gradients are neither necessary nor sufficient to the perception of a flat surface by showing that a surface with perspective and compression gradients but no density gradients reveals a flat receding surface, but that a density gradient alone, without perspective and compression, does not (see Marr, 1982, pp. 235–237). Yet the results of this study show that density and compression gradients together reveal flatness in a receding surface better than perspective alone for two out of four measures (see Table 1, Stimuli 4 and 5), even though perspective by itself is the single most important gradient.[4] This kind of quantitative analysis demonstrates the shortcomings of certain qualitative demonstrations. For example, necessity and sufficiency may not always be the firmest of grounds on which to build systematic theories of perception. The method of strong inference (Platt, 1964) can be too strong, losing important subtlety in phenomena and giving mis-

---

[4] It is the preference scaling values, but not those for dissimilarities, that show this effect. Notice that it goes against the notion that 65% of the variance is accounted for by perspective and that 35% is accounted for by compression and density combined. Our only response to this conundrum is that these comparisons have comparatively high local stress (discrepancy) around them, suggesting a small violation of additivity of gradient information. The raw data in both cases show that Stimulus 4 (with perspective only) was preferred over Stimulus 5 (with compression and density) in 65% and 70% of all direct comparisons in the flat/regular and flat/irregular tasks, respectively.

impressions of the combinatorics of information for visual perception. Of course, given that we have used a single tilt, a single viewing window, and one-dimensional gradients, we are subject to the same criticism. The point remains valid, however, and we will be only too pleased when better and more accurate assessments than ours are made.

Let us now turn to why it is that perspective is the salient gradient for the perception of flat surfaces, and compression that for curved surfaces.

Perspective, Flatness, and an Invariant

There are probably many ways in which a human perceiver or a computer vision system might use gradient information to judge the flatness of a surface. Because our data show that human perceivers appear to use the perspective gradient as most reliable in this domain and because the other important gradient, density, is under normal circumstances computationally related to perspective (compare Equation 1 with the last part of Equation 3 under Assumption 3), our tack is to try to determine how a perceiver may determine flatness from perspective.

We make the assumption that the visual system is extracting some information from the visual display, and that that information is invariant. Although the notion of invariants for perception is not without problems (see Cutting, 1983), invariants can provide anchors of information for perceptual systems that can guide perceptual process. We also explore the ecological validity (Brunswik, 1956) of these sources of information in hopes of providing further constraints.

Immediately, however, one should recognize that gradients cannot, by themselves, be invariants: Things that grade, or change, with visual angle cannot also be constant, or invariant. Instead, the particular variant in mind, the perspective gradient, is assumed to be associated with some other specifiable information in the optic array that is invariant, or at least roughly so. We assume further that the invariant information in the optic array is generated by some uniform property of the textured surface under scrutiny. The latter assumption is essentially the same as we made earlier: Texture elements are roughly of the same absolute size. This assumption then feeds into the former: Since size—or in our case the horizontally measured visual angle—will grade with distance, we can use invariant relations of size and distance in the analysis of flatness.

The invariant that we are looking for is an information source that specifies flatness. One way to tackle this problem is to consider eye height, or e in the equations found at the beginning of this article, as an invariant of the situation where an observer finds him- or herself in relation to a plane. Eye height, or the distance to the eye from the plane as measured along a line orthogonal to that plane, is constant in relation to all points on the surface of the plane. If we can specify an invariant of eye height whose basis is in the variable perspective gradient combined with some other variable information, then we may have a plausible invariant that perceivers could use.

Let us begin with a modification of the perspective gradient from Equation 1:

$$P = 2 \cdot \arctan[k/(d^2 + e^2)^{1/2}], \qquad (4)$$

where we are no longer directly interested in the radius of a given texture t, but instead substitute a constant k for t. (The assumption of roughly equal sized textures is still made.) Let us next define an entity mentioned previously, optic α, which is the angle from any given texture on the plane to the eye and then to the point on the plane just beneath the eye. In this manner, the optic angle α for the horizon is virtually 90°, and that for any texture at one's feet is near 0°. Thus, α is essentially the information sometimes known as *height in the picture plane*. The value of α depends on two parameters already used before, distance from one's feet (or the point directly below the station point), and eye height:

$$\alpha = \arctan(d/e). \qquad (5)$$

Rearranging Equation 5 to express these relations in terms of distance, we find that $d = e \cdot \tan \alpha$. Substituting for d in Equation 4, and again rearranging the terms, we get the following:

$$\tan(P/2) = k/(e^2 \cdot \tan^2 \alpha + e^2)^{1/2}. \qquad (6)$$

Solving Equation 6 for eye height, we get the following:

$$e = k/[\tan^2(P/2) \cdot \tan^2 \alpha + \tan^2(P/2)]^{1/2}. \qquad (7)$$

At this point, we need to simplify this cumbersome expression. First, because the horizontal visual angle P of any given texture will not be very large, rarely larger than a few degrees of visual angle, $\tan^2(P/2)$ will be very small. For example, $\tan 2° = .03$, and $\tan^2 2° = .001$. Thus, the value of $\tan^2(P/2)$ typically will be insignificant, and could be omitted from the last part of Equation 7. With its removal, the squares and square root operations cancel. Second, because for small angles a tangent transformation is very nearly a linear transformation, we can consider $\tan(P/2)$ to be equivalent to P times a constant. Accumulating all the constants (k, the divisor of P, and the constant of the tangent transformation for small angle) into a new constant c, we find the following:

$$e \approx c/(P \cdot \tan \alpha). \qquad (8)$$

In this manner, eye height is very nearly a straightforward multiplicative function of the perspective gradient P for any texture at any point in the visual field and $\tan \alpha$, where α is the optic angle to the texture associated with the particular P value. The results of Experiment 1, we believe, demonstrate that the perspective gradient and its P values are used by perceivers, and it is fairly obvious that viewers ought to be aware of α, or where within the field of view they happen to be looking with respect to their feet or to the horizon. For any given plane and the relation of a station point to it, $P \cdot \tan \alpha$ is very nearly constant and could serve as (nearly) invariant information on which judgments about the flatness of a surface could be made. Now, this is not to say that we think eye height is (directly) perceived; our only point is that invariant eye height is a corollary to the flatness of a plane, and that an invariant e could be information that the perceptual system uses. What the individual sees, and what viewers report seeing in our experiment, is a flat surface.

The force of this exercise is to demonstrate how an invariant that generates the perspective gradient may be useful in judging flatness. Of course, one can proceed through the same exercise for the other two gradients. In a sketchy way, let us consider density first. Because of the similarity between Equations 1 and 3, the same general end will be reached as in Equations 4–8, although not quite so simply. To make this type of analysis work for density one needs to exchange the assumption of roughly uniform texture size for one of roughly uniform spacing. In many situations, such as looking at trees on a savannah or at rocks on a desert, the latter assumption may be less warranted than the former. This fact, and perhaps the slightly increased cumbersomeness of the end expression (analogous to Equation 8), may contribute to the density gradient weighing less in our perceptual assay. Another possibility is that the perspective gradient, as shown in the middle panel of Figure 3, is very nearly linear for optic angles greater than 45°, whereas density is nearly exponential. This fact may make perspective information much more trustworthy than density as a perceptual anchor.

Consider next the compression gradient. Because of the manner in which compression is determined in Equation 2, distance d quickly falls out of the equation when rearrangements are made, and there is then no way to substitute α, or the general height in the plane of a texture, back into the equation. Thus, one cannot easily fix eye height to compression values within the visual field. It would seem that without this anchor, one would be at a disadvantage, if not a loss, to determine e, although perhaps some other specification of flatness could be made. Another problem with the use of compression is that it assumes that textures lie flat on the plane. Rocks, trees, and animals are usually clear violations of this assumption, perhaps rendering compression information somewhat less useful for this purpose. We return to the role of compression in perception of surfaces later.

This, of course, has been only a formal analysis of why perspective may be so important in the perception of flat surfaces. We regard it as plausible, but not necessarily definitive. More empirical work is needed. The essence of our argument is that the perspective gradient of textures on a flat surface is specified by uniform eye height, which in turn is cospecified by the flatness of the surface. This may seem circular in logical form, but it is not. It boils down to the idea that the optical horizontal extent of textures or objects (of roughly uniform size), in conjunction with information about their height in the field, is sufficient information to specify flatness. This

characterization is an unfamiliar form of a familiar pictorial cue, linear perspective.

Notes on Linear Perspective

Linear perspective, particularly as revealed in the familiar example of railroad tracks disappearing into the distance, is measured exactly this way. That is, under conditions of no tilt, the width between the tracks is directly related to height in the picture plane (or α). Because the bed on which the tracks are laid is flat, e is constant, and the tracks themselves project linearly.

One other fact about linear perspective should be considered. Linear perspective occurs under conditions where there is no variation in absolute size of elements and where there is also no variation in absolute density. In the railroad track example, the rail textures remain exactly the same absolute size as they recede from view, and the distance between the middle of the rails (hence, density) is also exactly the same. In the present experiments we have decoupled the two, generating no variation in size but statistical variation in density for surfaces with regular texture elements, and generating independent variation in size and density for surfaces with irregular texture elements. In both situations perspective and density accounted for a combined 90% of the variance or better in scale values across all measures, suggesting further independence of these gradients. That is, regularity in the perspective gradient accounted for no more of the variance than did perspective under statistical variation (regular vs. irregular-shaped textures). We anticipate that in viewing situations with linear perspective, the coupled perspective and density gradients would also account for about 90% of the variance. Such a result would be consistent with Attneave and Olson (1966), Braunstein (1976), and Gillam (1970).

Compression, Curvature, and a Variant

As with judgments of flatness, there are probably many ways in which a human or computer vision system may make judgments of curvature. Because our data demand that we consider compression gradients alone, we will do just that. But first consider some arguments why perspective and density gradients may not suffice for curvature judgments.

The perspective gradient essentially registers distance. In three-dimensional settings, vertical deviations from flatness are relatively small when compared with horizontal extent. Consider for a moment a plastic environment with a hill that can vary in its height: Potential changes in the overall distance from the eye to the peak of this particular plastic hill are very small, due to potential variations in the height of the hill. These Pythagorean relations are captured as $[(e - h)^2 + d^2]^{1/2}$, where the new term h is the hill height. Since both e and h will usually be small compared with d, the overall distance from eye to peak will be determined primarily by the horizontal distance factor d. That this is true is shown by comparing the perspective gradients in the middle and lower panels of Figure 3. The fact that the perspective gradient changes little when a surface becomes nonflat makes it an unreliable source of information in registration about deviations from flatness.

The density gradient also measures distance, in part, and it does this through spacing of texture elements. Scrutiny of Equations 1 and 3 shows that density and perspective are related if t and n are related, but comparison of Equations 2 and 3 shows that the same is true for density and compression. It is this latter relation that contributes to the sharp inflection shown in the bottom panel of Figure 3 for the density gradient. One may think that the peak in this gradient would be a salient local cue for the registration of curvature. However, it may be that this peak is mostly a function of the continuous representation of density as shown in the figure; in the actual stimuli an interpolated density function may not be nearly so articulated. That is, given that there was only an average of 13 "rows" of textures within the 15° window that comprised the projection of each stimulus, there is on average less than 1 row per degree. Because the peak, as it would reveal a nonmonotonicity in the density function, is less than 2° wide, there may be too few textures within this region to register the much increased local density. Thus, statistical variations that are present in these stimuli and that would be likely in the real world may render density a less efficacious source of information about curvature.

This leaves compression. Because there is no absolute shape that is constant for all hills, we will not look for invariant information about curvature. Instead, because the compression gradient changes broadly and nonmonotonically for the curved-surface stimuli, as shown for the C function in the bottom panel of Figure 3, aspects of this variable function will be entertained. The question is: What is it about these changes that makes compression a salient source of information about curvature? The answer is that compression is part of the crucial information that signals to the observer the relative orientation of the texture element with respect to the observer. Optical θ, or slant, is the angle between the line of sight and the line passing through the point of regard that is orthogonal to the surface, or the surface normal (see Flock, 1964; Stevens, 1981b; Witkin, 1981). It is defined by the following relation:

$$\theta = \arccos(C/P), \qquad (9)$$

where C and P are the Equation 2 and Equation 1 gradients for any texture element at any optic angle α. As mentioned earlier, C/P is the compression gradient, properly measured. Optical θ for both flat and curved surfaces, then, is a function of compression, and it is the ratio of the functions shown in the middle and bottom panels of Figure 3.

Experiment 2: Assessment of Thresholds for Curvature Perception

The compression gradient is shown in Figure 4 for several types of surfaces. Function 1 represents the gradient for the curved surfaces used in Experiment 1, and Function 4 represents that for flat surfaces. Function 2 is the gradient for a hill of the same shape as Function 1, but only half the height. Function 3 is for a hill only one quarter of the original hill height. For hill heights only a little greater than that used in Experiment 1, the ratio dips below zero, indicating that occlusions have occurred. Since P remains essentially linear, most of the inflection in these functions is due to variations in C.
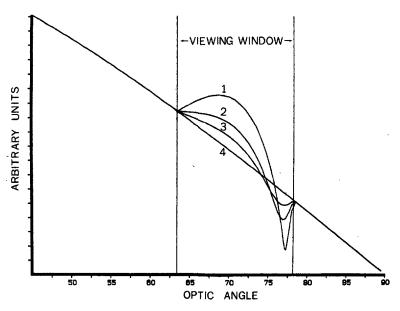

Figure 4. The compression gradients (C/P) for those surfaces shown in the top panel of Figure 3, and for two other surfaces used in Experiment 2. (Function 1 indicates the gradient for the curved surface used in the experiment; Function 2 is for a surface of the same shape but half the height; Function 3 for one-quarter the height of that for Function 1; and Function 4 for a flat surface. The compression gradient is the information available to perceivers to determine slant, the angle between the surface normal and the line of sight. Observers can judge that hills corresponding to Functions 1 and 2 are curved, but they cannot for that of Function 3.)

One can see that Function 1 is highly articulated. A large inflection occurs nearly seven-eighths up the viewing window, and the location of this inflection can be discerned in Figure 2 for Stimuli 2, 5, 6, and 8. Indeed, all viewers in the experiment said that they were particularly aware of this region in their judgments of the curved surfaces. Going back to Figure 4, one notices that the smaller the hill, the less articulated the function.

We strongly suspected that the nonmonotonicity of the compression function was the source of information used by perceivers in their judgments of curved surfaces. In order to determine if there was merit in this idea, we ran four observers through a variant of the curved-surface task employed in the recently described experiment. Using Stimulus Types 7 and 8 from Figure 2 in both regular and irregular versions, we had observers make pairwise comparisons as to which stimulus looked more like a curved surface. We started the observers with hills identical to those shown in the figure, with maximum elevations of 0.14 eye heights. All performed perfectly at this hill altitude, preferring Stimulus 8 over 7. We then halved the altitude (0.07 eye heights), which generated the compression function shown as Function 2 of Figure 4. Overall performance fell to 73%, but was still reliably above chance. Halving the hill height again (0.035 eye heights), generating Function 3, reduced performance to chance. Again, there were no strong differences between regular and irregular textures on the surfaces; in fact, irregularly textured surfaces with hill height elevations of 0.07 eye heights were slightly more discernible than were their regularly textured counterparts. We find it remarkable that the visual system is so little perturbed by statistical variations found in the irregularly shaped textures.

We take these data, in connection with the functions shown in Figure 4, as indicating that observers determine the presence of curvatures in surfaces by discerning whether there are discernible nonmonotonicities in the compression gradient. Moreover, discernible nonmonotonicities are present when observing surfaces that have smooth curvatures whose maximum coplanarities are only 0.07 times the eye height, within a window that is between 2 and 5 times the eye height. Thus, in one static view a viewer about 1.67 m (5 ft 5 in.)

in height could discern a smooth curvature in a tiled floor of about 0.1 m (4 in.) in height, when that bump occurred throughout a distance of 3 to 7.7 m (10 to 25 ft). With less smooth bumps of more local extent but the same height, more striking changes in compression would occur, perhaps even generating occlusions. It seems likely that these would be even more discernible. At greater distances, because of the geometry of occlusions, such hill sizes may be even more detectable than they are between 2 and 5 eye heights. Add motion gradients and the ability to detect noncoplanarities would surely get much more acute.

In summary, then, compression (measured against perspective) determines the orientation of a particular patch of texture in relation to the line of sight of the viewer. This angle, called optical θ (Flock, 1964), is the much researched entity called slant. It appears that changes in slant, per se, are not so effective in surface judgments as are changes in the changes of slant. Slant for a flat surface cannot be relevant because we know that compression accounts for so little variance in these judgments. Thus, we believe that most of the literature on slant judgments has misapplied the phenomenally important aspect of slant: Slant does not concern the phenomenal impression of flat surfaces, it concerns the impression of curved surfaces.

Conjectures About Process

There are few inferences that we can make about perceptual process, but some logical considerations about process are in order. Perspective and compression, as we have construed them, are orthogonally measured gradients. Perspective is measured at right angles to the axis of tilt, and compression is the ratio of information along that axis as a function of perspective. Thus, the registration of tilt (Stevens, 1983) is logically prior to registration of gradient information shown useful in these studies.

Given roughly circular textures that project as ellipses on the image plane, the line along the minor axis of any texture element will suggest a local axis of tilt (Stevens, 1981b, 1983; Witkin, 1981). Marr (1982, p. 234) suggests that this information is extracted explicitly, and we agree. This axis, if general to

the optic array, sets up the coordinate system for measuring the three gradients. So long as all textures agree in this measure (and all did in the studies reported here) the entire plane will have the same tilt. Small perturbations in tilt by one texture in a field of others with identical tilts are probably quite salient, since viewers are quite good at determining the tilt of individual patches of texture (Stevens, 1983; see also Witkin, 1981). These perturbations indicate violations of global coordinates for perspective and compression measures, and among other things would immediately suggest the plane is not flat.

Following Marr (1982) and his students, we suggest that tilt is extracted locally from each texture and from all textures in parallel throughout the array. Global comparisons would then be made across the array. We suggest further that within any region of the optic array where a surface has generally uniform tilt (and such regions are not at all rare in the natural environment) comparisons among textures will be made for perspective and compression. Comparisons along the perspective gradient are relevant for determinations of flatness; comparisons along compression, for curvature. We make no claim that all the information in any gradient is used in making judgments about flatness or curvature, only that some of it must be used because it is these gradient relations that carry all the information about textures in our displays. We do suggest, however, that all the gradient information is likely to be registered in relatively early stages of visual processing. In other words, we suggest that gradients are there (in the stimulus and visual system) to be used.

It is certainly possible that other kinds of processes are used by the human visual system in surface perception. Marr (1982, pp. 234–236), for example, doubts the efficacy of gradients in this domain, suggesting that they are derived measures that may have little to do with perceptual process. We will return to this idea several times while addressing three other issues.

Three Other Considerations

We wish to address three other points: that of local versus global analysis of textures, that of mode of projection (orthographic vs. per-

spective), and that of the separability of gradients in the natural environment.

First, consider the issue of local versus global analyses. The issue here is often couched in terms of priority—which comes first? Whereas this may be an important issue in some situations (Navon, 1977; Pomerantz, 1983), we do not regard it as important here. What is important, we believe, is the level of analysis that is most pertinent to perceiver and percept. For machine vision, Marr (1982) and others (Stevens, 1981a; Witkin, 1981) seem to strongly prefer local approaches to global ones, and emphasize the utility of local analyses. That is, analysis begins on each local texture. The orientation of the texture is then discerned and the relations to other textures, computed. The relations to other textures are guided, in part, by the assumption that neighboring elements have smoothly changing differences between them (Marr, 1982, pp. 44, 49). This smoothness assumption, we believe, almost guarantees measurable (and even differentiable) gradients. Gradients, in our view and as suggested above, are the global information measures of surfaces. Since one needs textures before one can have gradients of textures, we have no qualm with the necessity of local operations before global ones. Our point is that certain relations among the local elements emerge after this analysis, that these relations are the gradients, and that the gradients are the phenomenally relevant level of analysis. Claims against the effectiveness of gradients (e.g., Marr, 1982, pp. 233–239) would appear to be based in assertions about priority in the local/global issue, in incomplete understanding of the richness in different kinds of gradients, and in the failure to realize that the smoothness assumption is precisely what can be tested by research on gradients.

Second, we believe that perspective (or polar) projection is absolutely essential to research on the visual perception of surfaces. We find it odd that most machine-vision approaches have used orthographic (or parallel) projection (Marr, 1982; Stevens, 1981a; Ullman, 1979; Witkin, 1981; see also Sedgwick, 1980, and Todd, 1982, for analyses of projective methods). Our objections to orthographic projection in general are twofold: (a) Most objects under our scrutiny in the real world are relatively near to us. Perspective projection cap-

tures variation in nearness; orthographic projection simulates infinite viewing distance, a situation approximated only when looking through telescopes at small, distant objects. Thus, orthographic projection is "unnatural." (b) Orthographic projection nullifies all three gradients. That is, it is unnatural in a particular way. Under orthographic projection, and under the assumption that textures are roughly the same absolute size, flatness, and spatial separation, the gradients cannot possibly be useful to a perceiver because they will not be present in the stimulus. Yet we believe that our results show that gradient information is crucial in judgments of flat surfaces receding in the distance.

To be sure, Ullman (1979, chap. 4) proposed a hybrid "polar–parallel" scheme, where local analyses are carried out in an orthographic system of coordinates and then combined later by comparing different axes of projection into a polar system. This scheme, according to Ullman (1979), is both computationally manageable and perceptually plausible. We have no qualm with either of these points, nor with the general scheme, but it does seem inelegant. Why not start with a polar scheme rather than adding the extra step of converting to it? Ullman (1979, p. 155) has three replies: He claims that a polar perspective system at initial stages of image analysis (a) is inherently susceptible to errors because of the relative smallness of perspective effects, (b) is usually computationally more complex than a parallel system, and (c) leads to performance superior to that usually attained by the human visual system. In response, we would say that (a) because the results of our experiments suggest that small projective effects perfusing large regions of the optic array can lead to large perceptual effects, redundancy can efficiently overcome measurement error in the image, (b) given our basic ignorance about the algorithms of the visual system and our increasing knowledge about their complexity, there is no real reason to feel constrained in our choice of computational algorithms in biological systems by the ease of implementation of them on digital computers (see Marr, 1982, chap. 1), and (c) assessments of what the visual system can do are always filtered through responses, intervening processes, and motivational concerns that can, and almost certainly do, degrade our

estimates of its ability. In other words, although we feel that Ullman's polar–parallel scheme is workable, a purely polar scheme may be equally workable for a complex biological machine, and the latter scheme is our esthetic preference. Esthetic judgments aside, there may be little empirical justification to choose between the two.

Third, some researchers may object to our orthogonal manipulation of gradients. The world, they may argue, does not present itself to viewers with noncorrelated sources of information as we have over the course of Experiment 1. We have two responses to this point. (a) Our data, in the sense that the scale solutions show additivity of gradient information, do not show any systematic interactions or synergies of information. Thus, viewers appear to treat the information separably. Of course, a counter to this response is that the demand character of the experimental setting and of the stimulus sets encouraged independence. Our parry, then, is simply to state that this is exactly what we hoped to show: Viewers can treat the various, independently specified gradients as independent. (b) More deeply, however, we note that the three gradients investigated here are based on independent assumptions about the real world. Some of these assumptions hold in some environments, but not in others, and the match across environments and assumptions seems far from perfect. Only in highly artificial environments—for example, when looking at tiled floors inside buildings—are these three gradients indivisibly yoked.

In natural environments the three gradients are, at best, moderately correlated. For example, consider the case when looking out a second-story window at a piazza crowded with sightseers, where each person is considered a texture element. The assumption of equal absolute size of people–textures seems likely to be more valid than the assumption of equal distances among people–textures, which in turn is more valid than the assumption that the people lie flat on the plane (which is patently false although they do stand on the flat plane). In contrast, consider the case when looking at the litter on the floor of an arena after a sporting event or a rock concert, where the pieces of litter are the texture elements. The assumption that the litter lies flat on the

plane seems likely to be more valid than either the assumption about equal spacing or that about equal size. It seems likely, however, that in the first case a flat surface is implicit, where the planar elements consist of people's heads above the piazza; and in the second case the flat surface is equally implicit, where the planar elements consist of the shreds of paper and plastic. Thus, we believe that the world is furnished in such a way that gradients are separable. Given the less-than-complete coupling among the three assumptions, and their corresponding gradients, separable sources of information may be necessary for systematic perception in the less-than-systematic environments that we find ourselves in every day.

Conclusions

Throughout this article we have concerned ourselves with three assumptions. The first assumption is that texture elements on a surface are of roughly the same size. By this we mean that although there may be textures differing in size as boulders and blades of grass, most all boulders are of the same size and most all blades of grass are of the same size. We do not see that statistical variation is a problem, and to a remarkable degree the human visual system doesn't either: The results with randomly shaped (and thus differently sized) textures were no different than those with regularly shaped textures on either flat or curved surfaces. This size assumption is the basis for the regularity of the perspective gradient. For flat surfaces, two regularities fall out of these relations. First, the optically measured horizontal extent of any texture element is a function of distance; and second, distance is a function of optical angle α, or height in the plane. These relations, of course, are not perfectly invariant, but they are generally so and can serve as the information source that perceivers use about the flatness of a surface.

The second assumption is that texture elements lie flat or nearly flat on the plane. This assumption is probably justifiable in fewer situations than is the first assumption; as mentioned earlier, trees, animals, and many other objects that we find on surfaces are not functionally flat. Even grass yields a compression gradient inverse to that of flat textures: Looking down on the blades foreshortens them

more than looking at them a few yards away from one's feet. But absolute flatness against the surface need not always be attained, because the randomly shaped textures can be interpreted as objects, like rocks, which erupt out of the surface to somewhat different heights. Also, compressionlike gradients can be found elsewhere than on the surface of support. A flock of sheep seen from a distance reveals a compressionlike gradient as measured across the flanks and backs of each animal. The flatness assumption, of course, is the basis we use for compression gradients. As discussed earlier, it is also the basis of measures of slant. We believe that the data of the present study demonstrate that compression is not an important gradient for the perception of a flat surface, but that it is crucial for the perception of curvature. What observers appear to discern are changes in the changes of compression, or the second derivative. More particularly, nonmonotonicities in the compression gradient appear to be salient local sources of information about curvature. Unlike the use of perspective information, especially as it reveals the invariant of eye height and hence the flatness of a plane, the use of compression information appears to entail no invariant. Variants, particularly local inflections in the gradient, appear to be most important.

The third assumption is that texture elements have roughly uniform spacing. Because we were interested in perspective gradients, and not linear perspective, we chose to stagger the texture elements in all of our displays, making the statistical separation between elements uniform or nonuniform depending on our purposes. It seems to us that the assumption of even roughly equal distribution may be the most problematic of the three assumptions for most situations. The problem may not be so much that objects in the world are not roughly uniformly spaced; they may indeed be in more situations than not. But the uniformity of spacing may be out of scale to the purposes of the animal. Acacia trees may be uniformly spaced on a savannah, but this probably does not mean that their spacing would be a good source of information for flatness or curvature judgments. There are far too few of them. Perspective would be better for the former judgment, and treetop height for the latter. Relative densities may be either

too low as in this case, or too high in others as in the case of blades of grass (Flock, 1964), for density values to be perceptually relevant, or even perceptually discernible. The data of Experiment 1 appear to indicate that density gradients are of secondary importance, and probably useful only insofar as they reinforce perspective gradients, as in the case of linear perspective.

References

Attneave, F. (1972). Representation of physical space. In A. W. Melton & E. Martin (Eds.), *Coding processes in human memory* (pp. 283–308). New York: Halsted Press.

Attneave, F., & Olson, R. K. (1966). Inferences about visual mechanisms from monocular depth effects. *Psychonomic Science, 4,* 133–134.

Baird, J. C., & Noma, E. (1978). *Fundamentals of scaling and psychophysics.* New York: Wiley.

Bajcsy, R., & Lieberman, L. (1976). Texture gradients as a depth cue. *Computer Graphics and Image Processing, 5,* 52–67.

Beck, J., & Gibson, J. J. (1955). The relation of apparent shape to apparent slant in the perception of objects. *Journal of Experimental Psychology, 50,* 125–133.

Braunstein, M. L. (1976). *Depth perception through motion.* New York: Academic Press.

Braunstein, M. L., & Payne, J. W. (1969). Perspective and form ratio as determinants of relative slant judgments. *Journal of Experimental Psychology, 81,* 584–590.

Brunswik, E. (1956). *Perception and the representative design of psychological experiments.* Berkeley: University of California Press.

Clark, W. C., Smith, A. H., & Rabe, A. (1955). Interaction of surface texture, outline gradient, and ground in the perception of slant. *Canadian Journal of Psychology, 9,* 247–253.

Cutting, J. E. (1983). Four assumptions about invariants. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 310–317.

Epstein, W., Bontrager, H., & Park, J. (1962). The induction of nonveridical slant and the perception of shape. *Journal of Experimental Psychology, 63,* 472–479.

Flock, H. R. (1964). A possible basis for monocular slant perception. *Psychological Review, 71,* 380–381.

Flock, H. R., Graves, D., Tenney, J., & Stephenson, B. (1967). Slant judgments of single rectangles at a slant. *Psychonomic Science, 7*(2), 57–58.

Freeman, R. B. (1966). Absolute threshold for visual slant: The effect of stimulus size and retinal perspective. *Journal of Experimental Psychology, 71,* 170–176.

Gibson, J. J. (1950). *The perception of the visual world.* Boston: Houghton Mifflin.

Gibson, J. J. (1979). *The ecological approach to visual perception.* Boston: Houghton Mifflin.

Gibson, J. J., & Carel, W. (1952). Does motion perspective independently produce the impression of a receding surface? *Journal of Experimental Psychology, 44,* 16–18.

Gillam, B. (1970). Judgments of slant on the basis of foreshortening. *Scandinavian Journal of Psychology, 11,* 31–34.

Hochberg, J. (1982). How big is a stimulus? In J. Beck (Ed.), *Organization and representation in perception* (pp. 191–218). Hillsdale, NJ: Erlbaum.

Horn, B. K. P. (1975). Obtaining shape from shading information. In P. H. Winston (Ed.), *The psychology of computer vision* (pp. 115–166). New York: McGraw-Hill.

Kruskal, J. B. (1968). *How to use M-D-SCAL, a program to do multidimensional scaling and multidimensional unfolding.* Unpublished manuscript. Bell Telephone Laboratories, Murray Hill, NJ.

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

McArthur, D. J. (1982). Computer vision and perceptual psychology. *Psychological Bulletin, 92,* 283–309.

Navon, D. (1977). Forest before trees: The precedence of global features in visual perception. *Cognitive Psychology, 9,* 353–383.

Olson, R. K. (1974). Slant judgments from static and rotating trapezoids correspond to rules of projective geometry. *Perception & Psychophysics, 15,* 509–516.

Perrone, J. A. (1980). Slant underestimation: A model based on the size of the viewing aperture. *Perception, 9,* 285–302.

Phillips, R. J. (1970). Stationary visual texture and the estimation of slant angle. *Quarterly Journal of Experimental Psychology, 22,* 389–397.

Platt, J. R. (1964). Strong inference. *Science, 146,* 347–353.

Pomerantz, J. R. (1983). Global and local precedence: Selective attention in form and motion perception. *Journal of Experimental Psychology: General, 112,* 516–540.

Rosinski, R. R., & Levine, N. P. (1976). Texture gradient effectiveness in the perception of surface slant. *Journal of Experimental Child Psychology, 22,* 261–271.

Sedgwick, H. A. (1980). The geometry of spatial layout in pictorial representation. In M. Hagen (Ed.), *The perception of pictures,* (Vol. 1, pp. 33–90). New York: Academic Press.

Shepard, R. N. (1974). Representation of structure in similarity data: Problems and prospects. *Psychometrika, 39,* 373–421.

Stevens, K. A. (1981a). The information content of texture gradients. *Biological Cybernetics, 42,* 95–105.

Stevens, K. A. (1981b). The visual interpretation of surface contours. *Artificial Intelligence, 17,* 47–73.

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

Todd, J. T. (1982). Visual information about rigid and nonrigid motion. *Journal of Experimental Psychology: Human Perception and Performance, 8,* 238–252.

Todd, J. T., & Mingolla, E. (1983). The perception of surface curvature and direction of illumination from patterns of shading. *Journal of Experimental Psychology: Human Perception and Performance, 9,* 583–595.

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

Vickers, D. (1971). Perceptual economy and the impression of visual depth. *Perception & Psychophysics, 10,* 23–27.

Witkin, A. P. (1981). Recovering surface shape and orientation from texture. *Artificial Intelligence, 17,* 17–45.