The effects of viewing angle, camera angle, and sign of surface curvature on the perception of three-dimensional shape from texture

James T. Todd Department of Psychology, The Ohio State University, Columbus, OH, USA

Lore Thaler Department of Psychology, The Ohio State University, Columbus, OH, USA

Tjeerd M. H. Dijkstra Department of Information and Knowledge Systems, Radboud University, Nijmegen, The Netherlands

Jan J. Koenderink Department Physics of Man, Helmholtz Institute, Utrecht University, Utrecht, The Netherlands

Astrid M. L. Kappers Department Physics of Man, Helmholtz Institute, Utrecht University, Utrecht, The Netherlands

Computational models for determining three-dimensional shape from texture based on local foreshortening or gradients of scaling are able to achieve accurate estimates of surface relief from an image when it is observed from the same visual angle with which it was photographed or rendered. These models produce conflicting predictions, however, when an image is viewed from a different visual angle. An experiment was performed to test these predictions, in which observers judged the apparent depth profiles of hyperbolic cylinders under a wide variety of conditions. The results reveal that the apparent patterns of relief from texture are systematically underestimated; convex surfaces appear to have greater depth than concave surfaces, large camera angles produce greater amounts of perceived depth than small camera angles, and the apparent depth-to-width ratio for a given image of a surface is greater for small viewing angles than for large viewing angles. Because these results are incompatible with all existing computational models, a new model is presented based on scaling contrast that can successfully account for all aspects of the data.

Keywords: shape from texture, three-dimensional shape, shape perception, vision, texture gradients, scaling, foreshortening, viewing geometry, scene geometry

Citation: Todd, J. T., Thaler, L., Dijkstra, T. M. H., Koenderink, J. J., & Kappers, A. M. L. (2007). The effects of viewing angle, camera angle, and sign of surface curvature on the perception of three-dimensional shape from texture. *Journal of Vision,* 7(12):9, 1–16, http://journalofvision.org/7/12/9/, doi:10.1167/7.12.9.

Introduction

Among all of the different sources of visual information about three-dimensional shape, gradients of texture are perhaps the least understood. Consider, for example, the image in Figure 1 that depicts a planar surface slanted in depth, which is covered with a random pattern of circular polka dots. Because of the effects of perspective, the optical projections of these polka dots have variable sizes and shapes that are determined by their relative distances and orientations with respect to the point of observation. It is the overall pattern of these systematic variations that somehow provides sufficient information to determine the apparent planarity and slant of the depicted surface.

Since gradients of texture were first identified by James Gibson (1987, 1950, 1979), numerous computational models have been developed in an effort to exploit this information for the estimation the three-dimensional surface structure from monocular visual images. One popular approach for estimating local slant is based on an assumption that variations in reflectance on a surface are statistically isotropic. In the special case of polka dot textures as shown in Figure 1, the optical slant (σ) at the center of each texture element can be determined by the following equation:

$$\cos(\sigma) = \frac{\omega}{\lambda}, \quad (1)$$

where λ and ω are the major and minor axes of a texture element's optical projection. We will refer to this

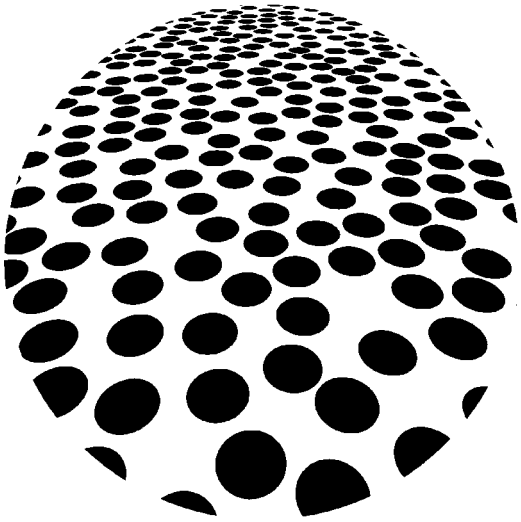

Figure 1. An image of a planar surface at a 50° slant with a polka dot texture.

approach as the analysis of texture foreshortening because that is the term that is typically used in the literature to describe the ratio ω/λ. Similar computations can also be performed for less regular isotropic textures from the distribution of edge orientations in each local image region (Aloimonos, 1988; Blake & Marinos, 1990; Blostein & Ahuja, 1989; Marinos & Blake, 1990; Witkin, 1981) or from the relative anisotropy of their local amplitude spectra (Bajcsi & Lieberman, 1976; Brown & Shvayster, 1990; Krumm & Shafer, 1992; Sakai & Finkel, 1994; Super & Bovik, 1995).

An alternative approach that comes closer to Gibson's original conception is to estimate surface slant by measuring the changes of optical texture across different local neighborhoods of an image, based on an assumption that the texture on a physical surface is statistically homogeneous. As was first demonstrated by Purdy (1958), the optical slant (σ) in a given local region can be determined by the following equation

$$\tan(\sigma) = \frac{2(\lambda_1 - \lambda_2)}{\delta(\lambda_1 + \lambda_2)}, \quad (2)$$

where δ is the projected distance between neighboring optical texture elements in the direction that slant is being estimated, and λ_1 and λ_2 are the projected lengths of those texture elements in a perpendicular direction (see Figure 2). In the limit of an infinitesimally small δ, the right side of Equation 2 equals the normalized depth gradient (Purdy, 1958, Equation 14; Gårding, 1992, Equation 33). Similar computations can also be performed on less regular textures from the affine correlations between the amplitude spectra in neighboring image regions (Clerc & Mallat, 2002; Malik & Rosenholtz, 1994, 1997) or from the systematic changes in the distributions of edges (Gårding, 1992, 1993).

When considering these alternative procedures for estimating slant from texture, it is important to recognize that Equations 1 and 2 are only valid when their optical variables are defined as visual angles rather than distances in the image plane. As a consequence of this, these methods can only produce accurate estimates of local surface slant when an image is viewed at the same visual angle as the one with which it was photographed or rendered. It is also interesting to note in this regard that the predicted effects of an inappropriate viewing angle are quite different depending upon which method is used to estimate slant. To demonstrate this more clearly, it is useful to consider the planar surface at a 50° slant that is depicted in Figure 1. This image was rendered with a 60° camera angle, but when viewed on a printed page at a comfortable distance, its visible angular extent will be much smaller than that (see Figure 3). Because reductions of image size increase the magnitudes of the optical texture gradients, they would also increase the estimated slant as computed from Equation 2. For example, if the image was observed from a 10° viewing angle, the information from optical texture gradients would specify that the depicted surface has an 83° slant relative to the frontoparallel plane. Reductions of image size have the opposite effect when slant is estimated from Equation 1 because optical foreshortening is reduced in peripheral regions relative to what would occur when an image is viewed from the correct visual angle. The estimated surface in that case would also be curved rather than planar.

To what extent does the perception of three-dimensional shape from texture conform with the predictions of either of these models? The research described in the present article was designed to address this issue. Human observers were asked to judge the apparent shapes of textured hyperbolic

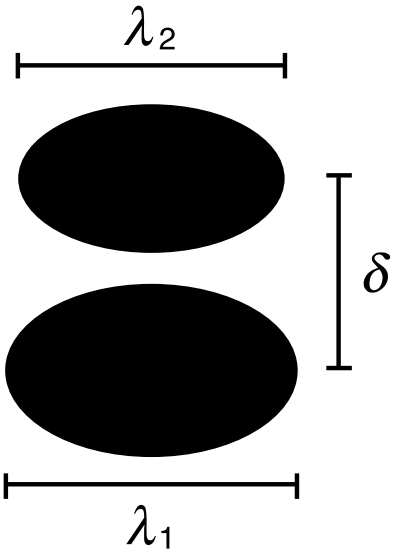

Figure 2. The variables used to estimate local optical slant from the gradient of texture scaling.

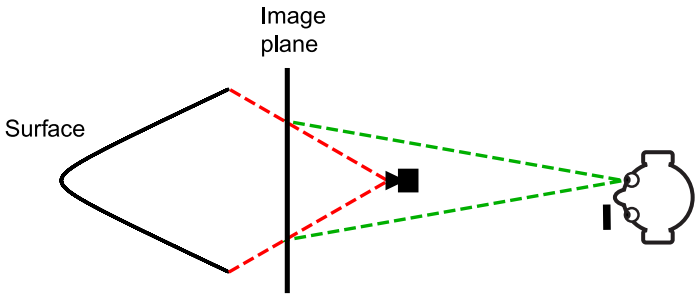

Figure 3. A schematic diagram of an image of a surface that is viewed with a different visual angle from the one at which it was photographed or rendered. The camera angle in this example is represented by dashed red lines, and the viewing angle is represented by dashed green lines.

cylinders from monocular images presented at different viewing angles that could be the same or different as the camera angles with which the images were rendered.

Methods

Subjects

Four of the authors participated in the experiment (A.K., J.K., J.T., and L.T.), and they all had normal or corrected-to-normal visual acuity.

Apparatus

The experiment was conducted using a Dell Dimension 8300 PC with an ATI Radeon 9800 PRO graphics card. Images could be viewed with a visual angle of either 20° or 60°. Those viewed at a 20° angle had horizontal and vertical extents of 30.0 cm and were presented on a standard CRT. Those that were viewed at a 60° angle had horizontal and vertical extents of 121.9 cm and were back projected onto a translucent display screen using an LCD projector. Both types of display had a spatial resolution of 1,280 × 1,024 pixels, and the stimulus images were presented within an area of 1,024 × 1,024 pixels. The displays were monocularly viewed with an eye patch, and a chin rest was used to constrain head movements.

Stimuli

There were 12 possible scene geometries that are shown in Figure 4. Each stimulus image depicted a vertically oriented convex or concave hyperbolic cylinder that was rendered with a camera angle of either 20° or 60°. The shapes of the surfaces were also varied, which was achieved by manipulating the angle (α) of their asymptotic lines relative to the frontoparallel plane (see Figure 5). Because prior research has shown that the apparent depth of a textured surface is attenuated by reduced camera angles or negative signs of curvature (Todd, Thaler, & Dijkstra, 2005), we increased the simulated depths in those conditions in an effort to ensure that none of the stimuli would appear completely flat. Thus, for the concave surfaces with 20° camera angles, the possible values of α were 55°, 60°, and 65°. For the concave surfaces with 60° camera angles and for the convex surfaces with 20° camera angles, the possible values of α were 50°, 55°, and 60°. Finally, for the convex surfaces with 60° camera angles, the possible values of α were 45°, 50°, and 55°. Note that all of the different combinations of camera angle and sign of curvature included a common asymptotic angle of 55° (see Figure 6).

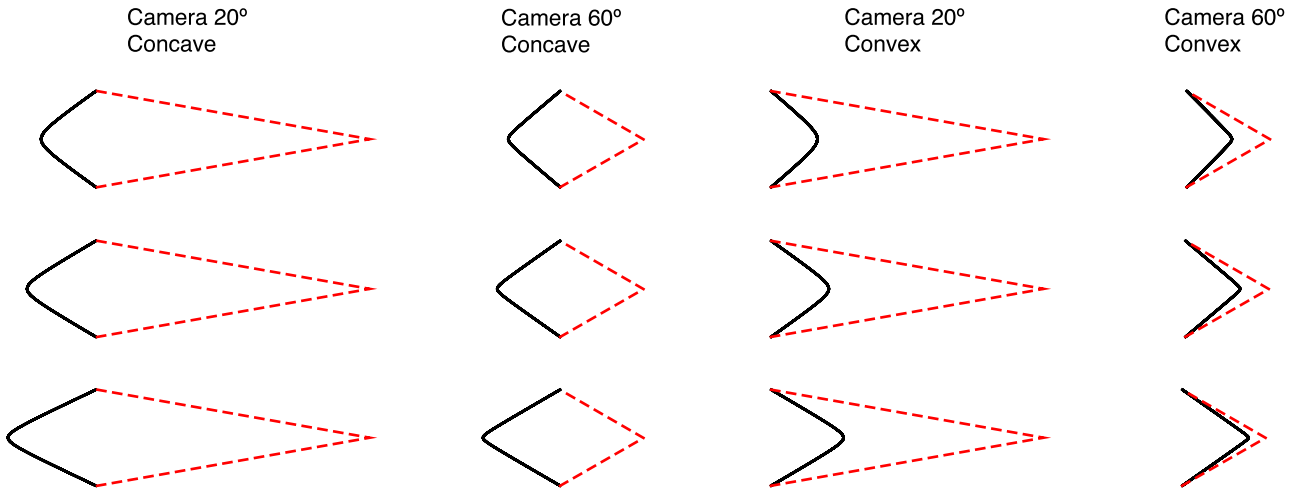

Figure 4. The different possible scene geometries that were used to generate the stimulus images employed in the present experiment. The black curves show horizontal cross-sections in depth of the depicted surfaces, and the dashed red lines show their corresponding camera angles.

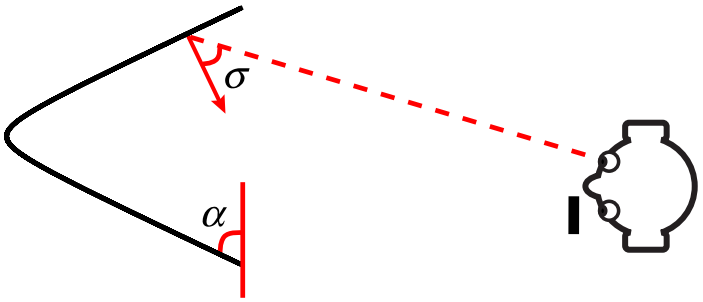

Figure 5. The asymptotic slant α of a hyperbolic cylinder relative to the frontoparallel plane. This should not be confused with the optical slant σ at any given surface point relative to the viewing direction.

The depicted surfaces were all presented with six possible textures that are shown in Figure 7. These included a pattern of horizontal lines, a pattern of noisy lines, a square grid of circular polka dots, a random pattern of polka dots, a random pattern of polka dots that varied in size, and a random pattern of ellipses with varying eccentricity. Three different versions were created for all of the random textures, which were randomly sampled as needed in the relevant conditions. For the textures with translational symmetry, a random phase was selected for each presentation. The textures were scaled on each surface so that the average projected element size would be the same in all conditions.

Procedure

On each trial, an image of a hyperbolic cylinder was presented on the main display screen directly in front of the observer. A second monitor was located off to the side of the main display that contained an adjustment figure defined by four parameters (P_1, P_2, P_3, and P_4), which observers could manipulate by adjusting four sliders with a hand held mouse to match the apparent cross-section in depth of the depicted surface (see Figure 8). Because some of the stimuli did not appear to have hyperbolic cross-sections, the adjustment space required four degrees of freedom to adequately match the observers' perceptions in all of the different conditions. The shape of the adjustment figure was defined by the following equation:

$$z = \pm\left(P_2\sqrt{1+\left(\frac{x\tan(P_1)}{P_2}\right)^2} - P_2 + \frac{P_3x^2}{2}\left(1 - e^{-.5\left(\frac{x}{P_4}\right)^2}\right)\right). \quad (3)$$

Note that the first and the second terms define a hyperbola, where P_1 controls the angle of the asymptotic lines and P_2 controls the curvature at its midpoint. The third term with parameters P_3 and P_4 was included so that the asymptotic lines of the hyperbola could be made to bow inward or

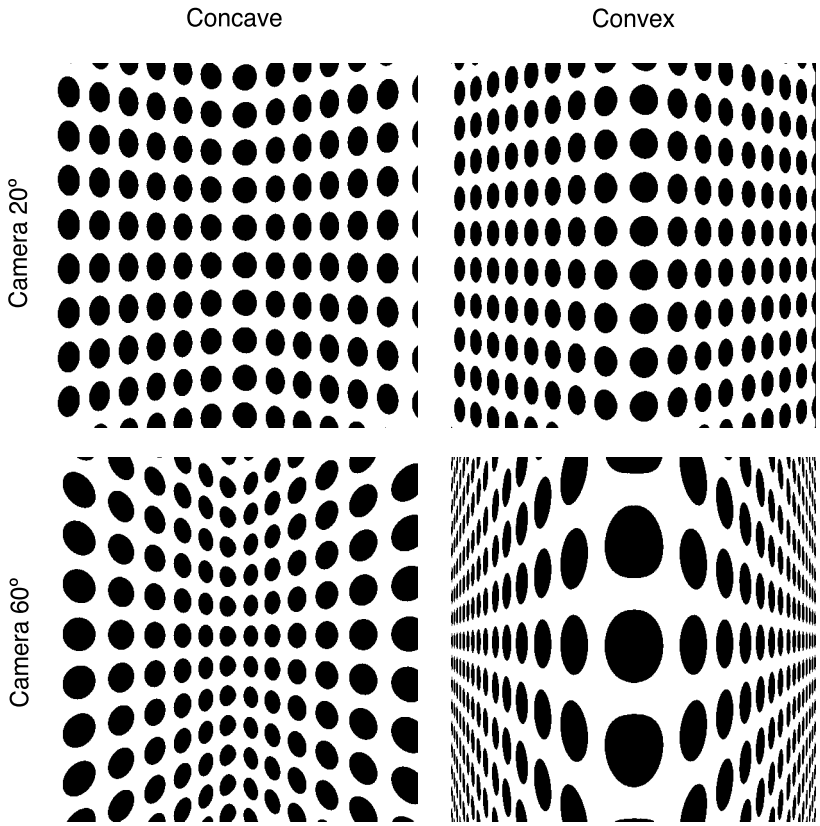

Figure 6. Example images of a concave and a convex surface that were rendered with different camera angles. The asymptotic slant of each depicted surface is 55°.

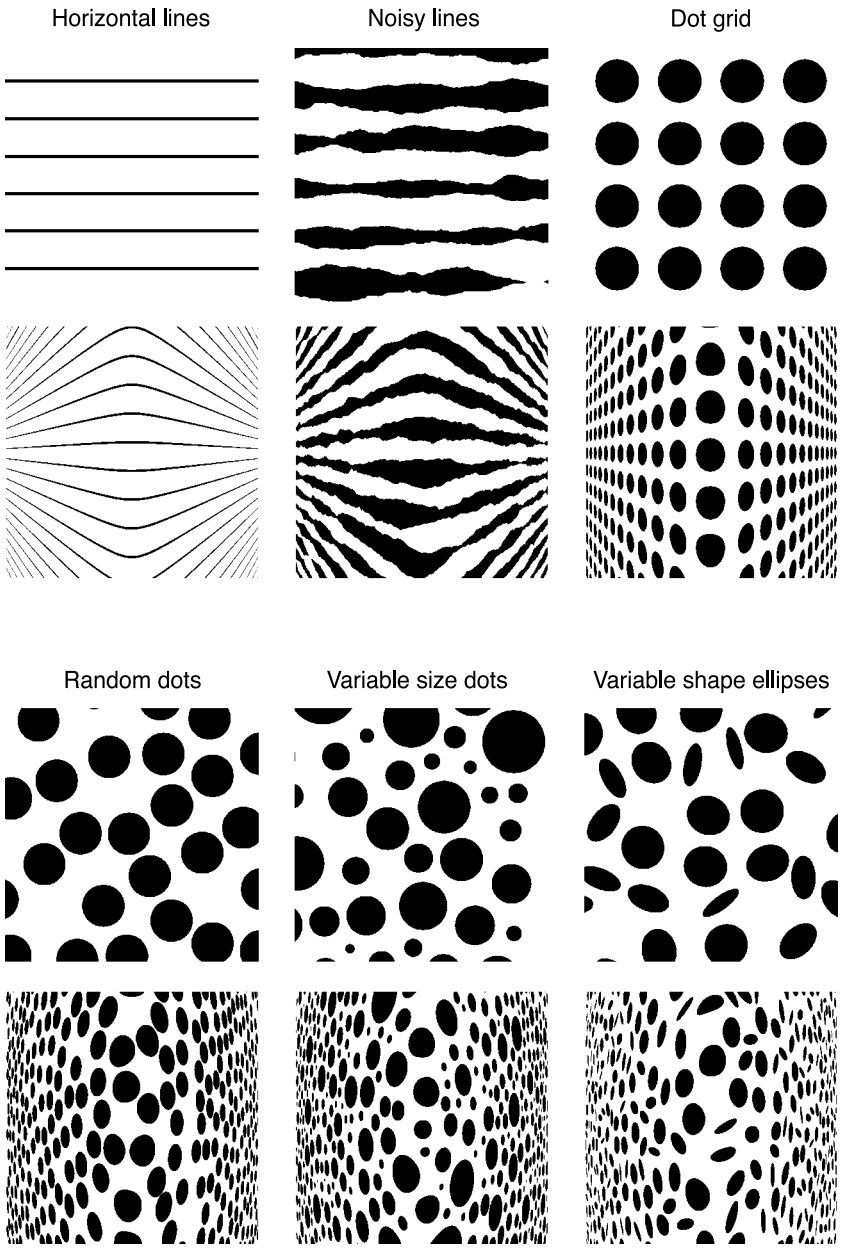

Figure 7. The six different textures used in the present experiment and an example stimulus that was generated from each one. The asymptotic slant of each depicted surface is 50°.

outward. The adjustment figure had a fixed width of 5.71 cm and had a height that varied from 0 to 13.2 cm.

At the beginning of each trial, the observers were required to click on one of three buttons to indicate whether the surface shown in the main display appeared convex, concave, or perfectly flat, and this response was used to constrain the orientation of the adjustment curve. They then adjusted the parameter P_1 with the appropriate slider to indicate the overall perceived relief of the depicted surface. Next, they adjusted the parameter P_2 to indicate the apparent curvature at its midpoint. Then, if necessary, they adjusted the parameters P_3 and P_4 to add some curvature to the asymptotic lines. Once observers were satisfied with their settings, they could move on to the next trial by clicking on a button that was labeled "next." All observers agreed that these response tasks were quite natural and that they had a high degree of confidence in their settings.

Design

To summarize the overall experimental design, there were 144 possible conditions: 6 different textures (horizontal lines, noisy lines, the dot grid, random dots, variable sized dots, and variable shaped ellipses) $\times$ 2 possible viewing angles (20° or 60°) $\times$ 2 possible camera angles (20° or 60°) $\times$ 2 signs of curvature (concave or

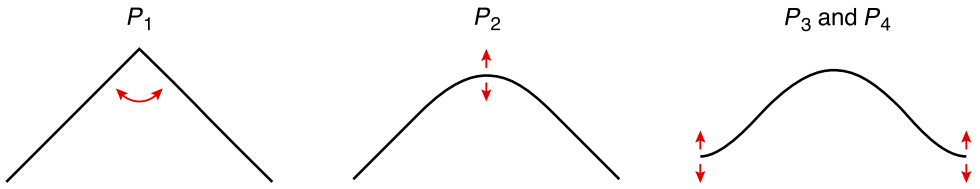

Figure 8. The parameters of the hyperbolic adjustment curve. P_1 controlled the angle of the asymptotic lines, P_2 controlled the curvature at its midpoint, and P_3 and P_4 caused the asymptotic lines to bow inward or outward.

convex) × 3 possible shapes for each combination of camera angle and sign of curvature. Within a given experimental session, the field of view remained fixed, and the 72 possible combinations of curvature, camera angle, shape, and texture were presented once each in a random sequence. Each observer participated in three separate sessions for each of the two possible fields of view.

Results

Overall, the observers were 98% accurate at determining the correct signs of curvature for these displays, and the small number of apparent depth reversals that did occur were restricted to those displays that depicted concave surfaces with a 20° camera angle. No reversals occurred for the horizontal line textures, and the percentage of reversals for the remaining textures ranged from 1% to 3%.[1]

With respect to the magnitudes of perceived depth, the results revealed that the variations in performance among the different experimental conditions had a high test–retest reliability. Comparisons of the judged depth-to-width ratios of a given observer across different experimental sessions had an average correlation of 0.95. There was also a high degree of consistency among different observers. The average correlation between each pair of observers was 0.96, although there were some individual differences in the overall perceptual gain of the settings (i.e., the judged depth-to-width ratio divided by the ground truth).

Although the observers were quite reliable, there were large systematic errors in the accuracy of their settings. Figure 9 shows the judged depth-to-width ratio as a function of the ground truth averaged over observers and textures for all of the different combinations of viewing angle, camera angle, and sign of surface curvature. Let us first consider the pattern of performance in those conditions where the displays were viewed at the same visual angle with which they were rendered. These are represented in Figure 9 by the black circles and red diamonds. It is important to keep in mind that when an image of a surface with isotropic texture is viewed from the correct visual angle, it is possible to obtain veridical estimations of the depicted depth-to-width ratio either from an analysis of the texture foreshortening or from an analysis of the local scaling gradients (see Equations 1 and 2). The results show clearly, however, that the observers' judgments were far from veridical. Note in Figure 9 that the depths of the depicted surfaces were systematically underestimated. In addition, the magnitude of perceived

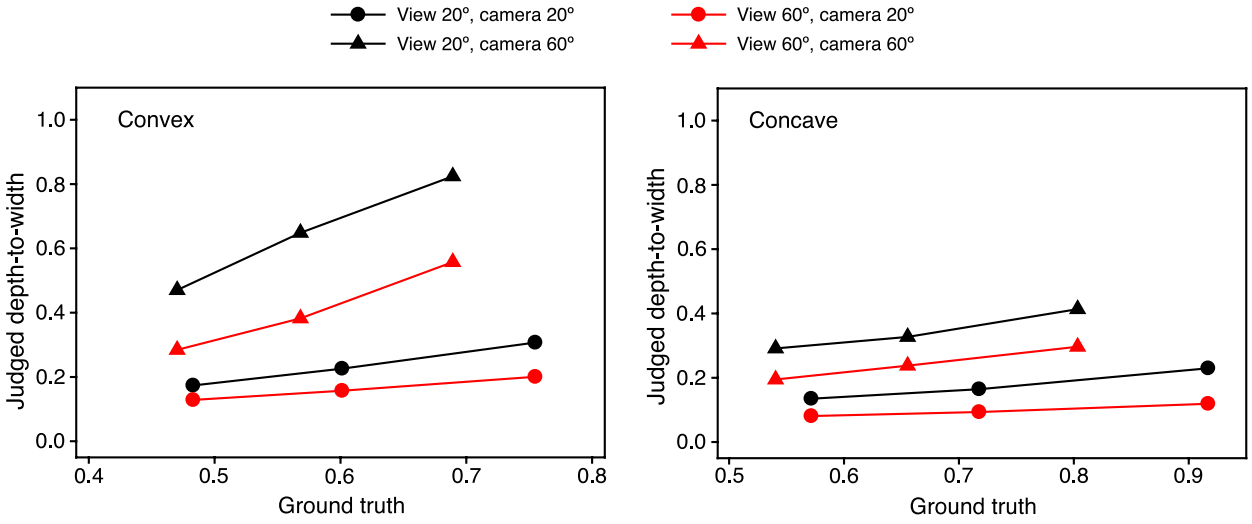

Figure 9. The judged depth-to-width ratio as a function of the ground truth averaged over observers and textures for all of the different combinations of viewing angle, camera angle, and sign of surface curvature.

depth was much larger for convex surfaces than for concave surfaces, and it was also much larger for the 60° camera angles than for the 20° camera angles (see also Tibau, Willems, Van Den Berg, & Wagemans, 2001; Todd et al., 2005). None of these effects would be predicted based on current computational analyses for determining three-dimensional shape from texture.

Another interesting finding that should be noted in Figure 9 is the differences in apparent depth that were obtained between 20° viewing angles (black symbols) and the 60° viewing angles (red symbols) when all other factors were held constant. If the three-dimensional shape of a depicted surface was estimated from local scaling gradients as described by Equation 2, then the judged depth-to-width ratio obtained with a 20° viewing angle should be 3.27 times larger than what would be obtained if the same image was viewed at a 60° angle (see Purdy, 1958). The opposite effect would occur if shape was computed from local texture foreshortening as described by Equation 1, although qualitative distortions in the estimated shape would be expected for reductions in image size if the optical texture elements become elongated in the direction of surface slant (e.g., see lower left panel of Figure 6). Which of these predictions is more consistent with human perception? The judged depth-to-width ratios in the present experiment were on average 1.55 times larger for the 20° viewing angles than for the 60° viewing angles. Thus, the direction of the effect is consistent with what would be expected from an analysis of local scaling gradients, but the magnitude is much smaller. Similar findings have also been reported by Backus and Saunders (2006) for the perceived three-dimensional slants obtained from linear perspective, although the magnitude of the effect they obtained is much smaller than the one reported here. This is most likely due to the large visual angles we employed because the optical distortions produced by variations of image size increase with retinal eccentricity.

Scaling contrast

The results of this study and the related findings of Todd et al. (2005) strongly suggest that observers' judgments of these displays cannot have been based on texture foreshortening or local gradients of texture scaling, but what other possible information might be available that could potentially explain how the magnitude of perceived depth from texture is influenced by the viewing angle, the camera angle, and the sign of surface curvature? To better understand these issues, it is useful to recall that for small visual angles, the projected length or scaling (S') of an optical texture element in radians can be closely approximated by the following equation:

$$S' \cong \frac{S}{D}, \tag{4}$$

where S is the diameter of a physical texture element and D is its distance from the point of observation (see Figure 10). One way of describing the global variations of scaling within an image is to incorporate a measure that will be referred to here as *scaling contrast,* as defined by the following equation:

$$\text{Scaling contrast} = \frac{S'_{\max} - S'_{\min}}{S'_{\max} + S'_{\min}} \cong \frac{D_{\max} - D_{\min}}{D_{\max} + D_{\min}}, \tag{5}$$

where $S'_{\max}$ and $S'_{\min}$ are the maximum and the minimum projected texture lengths, and $D_{\max}$ and $D_{\min}$ are the maximum and the minimum distances on a surface relative to the point of observation. Note that scaling contrast provides a reliable estimate of the surface depth contrast that is invariant over the size of the physical texture elements.

Another interesting property of this measure is that it is affected by the viewing angle of an image, the camera angle with which the image was rendered and the sign of surface curvature in much the same way as these variables influence observers' perceived depth. The left panel of Figure 11 shows the judged depth-to-width ratio averaged

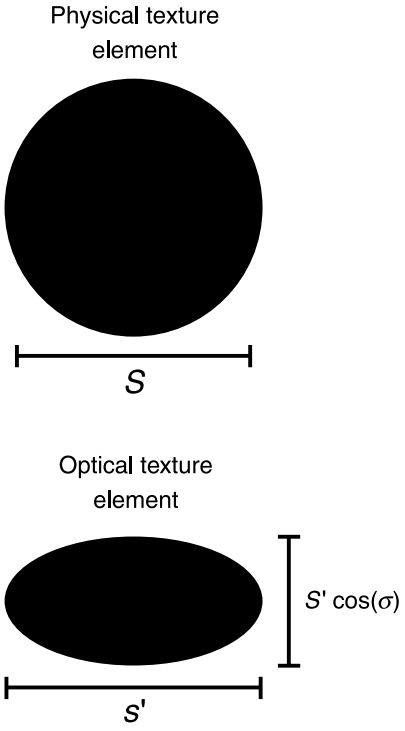

Figure 10. A circular texture element on an observed surface produces an optical projection with an elliptical shape. The major axis of this projected ellipse, called scaling, has a length S', which equals the diameter (S) of the physical texture element divided by its distance (D) from the point of observation. The length of the minor axis equals S' times the cosine of optical slant (σ). Scaling contrast measures the systematic variation of S' within an entire image.

over observers and textures as a function of scaling contrast. It is important to point out that the values of scaling contrast in this figure were not measured from the actual stimulus images. Rather, they were indirectly computed by calculating the projected lengths of idealized texture elements along a horizontal cross-section through the center of each image, and then calculating their visual angles based on the appropriate viewing geometry. Note that the measure provides a reasonably good fit to the data that accounts for 80% of the variance among the different experimental conditions. For purposes of comparison, the ground truth accounted for less than 1% of the variance. It is also clear from these data, however, that most of the residual variance was due to lower than expected perceived relief for the convex surfaces with a 60° viewing angle, which may have been caused by an inability of the visual system to process high frequency information in peripheral regions of the visual field (see also Todd et al., 2005). In an effort to correct for this, we added a free parameter to the model and determined that the best possible fits to the data could be obtained by ignoring any texture variations outside the central 45° of view for convex surfaces. The results obtained with this modification are presented in right panel of Figure 11. The modified measure provides a near perfect fit of the data that accounts for 97% of the variance in observers' judgments. This same measure also provides excellent fits for the individual textures (see Figure 12), although the overall perceptual gain was somewhat attenuated for the two line textures.[2]

Thus far, we have only considered relatively coarse scale measures of the judged surface shape (i.e., the depth-to-width ratio and the sign of surface curvature), but it is important to keep in mind that the observers' settings involved four adjustable parameters, which allowed a large range of possible depth profiles for any given depth-to-width ratio and sign of surface curvature. Figure 13 shows the average adjusted depth profiles collapsed over observers and textures (red curves), as well as the ground truth (black curves) for each depicted surface shape with each combination of viewing and camera angle. Note in particular how the apparent local curvatures were systematically distorted relative to the ground truth, especially for those displays with 60° camera angles. For the concave stimuli, the linear asymptotic lines of the surfaces often appeared to bow outward. Similarly, for the convex surfaces, the apparent curvature at the near point was systematically reduced relative to the ground truth, even when the judged depth-to-width ratio was close to veridical.

In an effort to model the judged depth profiles, we calculated the projected length (S'_ϕ) of an optical texture element that was centered at each visual angle (ϕ) along a horizontal cross-section through the center of each display. For the displays that were viewed at the correct visual angle, the pattern of projected lengths in an image can be used to determine the veridical depth-to-width ratio and shape of a horizontal surface cross-section using the following equations:

$$X_\phi = \frac{\sin(\phi)}{S'_\phi}, \tag{6}$$

$$Z_\phi = \frac{\cos(\phi)}{S'_\phi}, \tag{7}$$

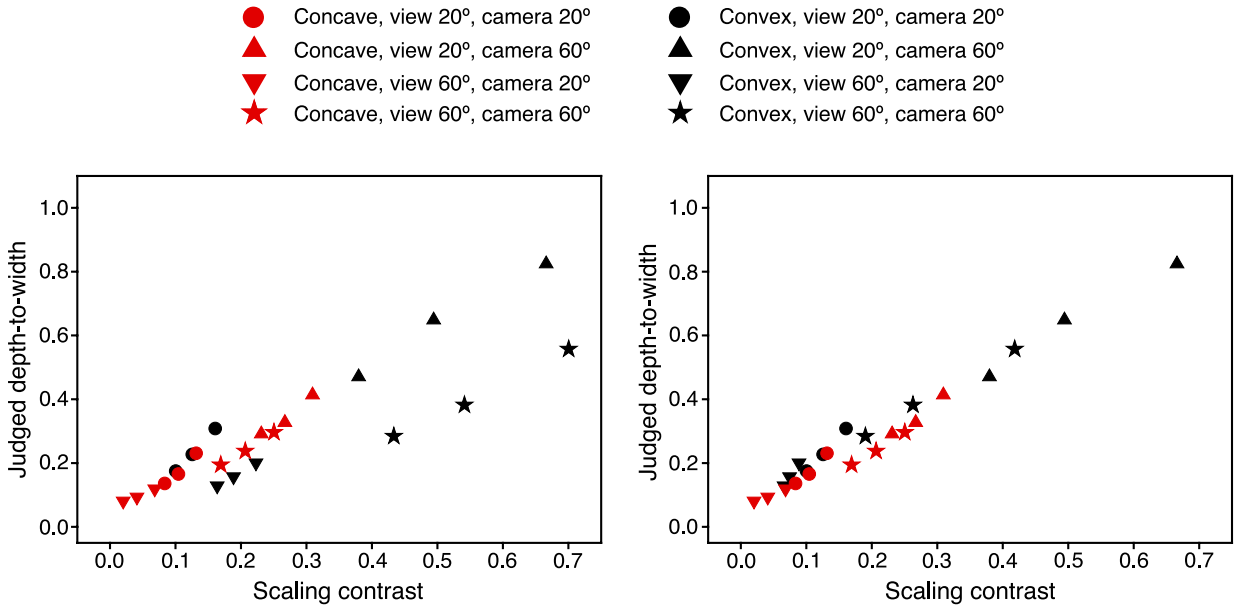

Figure 11. The judged depth-to-width ratios averaged over observers and textures as a function of scaling contrast. The values of scaling contrast in the right panel were restricted to the central 45° of view for convex surfaces, which produces a better fit to the data.

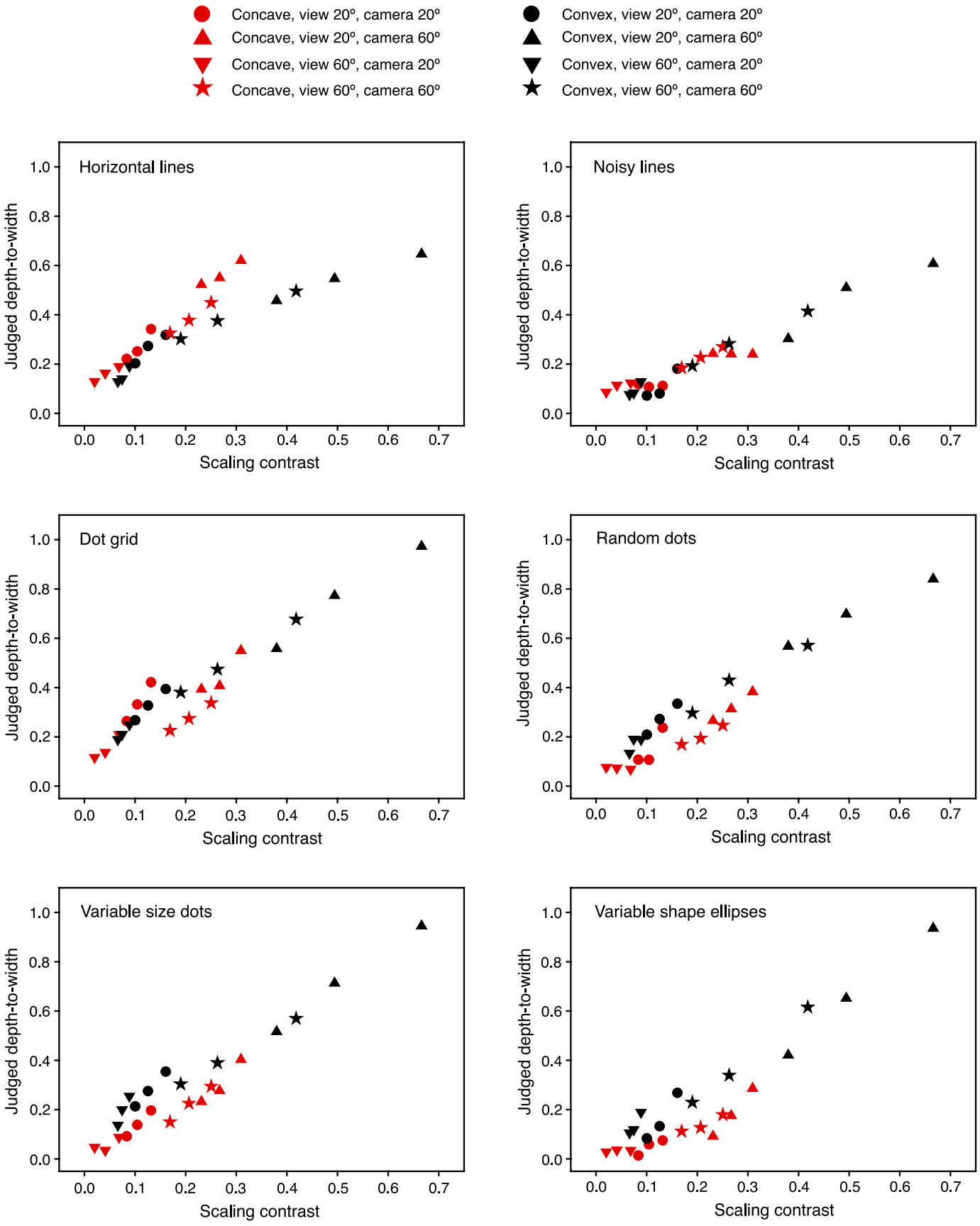

Figure 12. The judged depth to width ratios averaged over observers as a function of scaling contrast for each individual texture. The values of scaling contrast were restricted to the central 45° of view for convex surfaces.

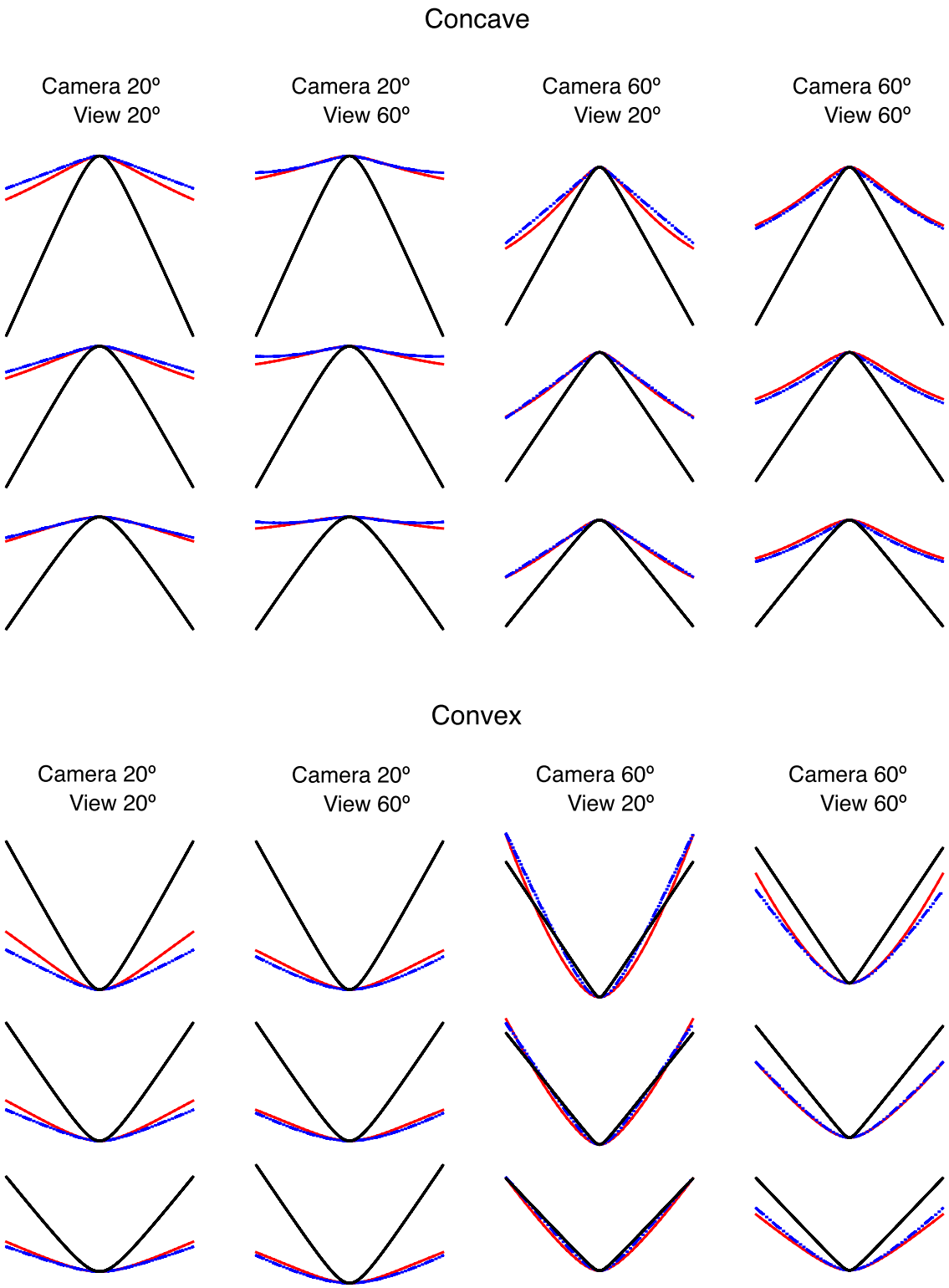

Figure 13. The judged depth profiles in each condition (red curves) averaged over observers and textures, the ground truth (black curves), and the estimated shapes from Equations 8 and 9 (dashed blue curves).

where X_ϕ defines the horizontal position of each point along the surface cross-section, and Z_ϕ defines its position in depth. It is clear from Figure 13, however, that observers did not adopt that strategy. We have tried several different procedures to model the judged shapes from the spatial variations of S'_ϕ, and the best fits to the data were obtained using the following equations:

$$X_\phi = \frac{\phi}{\phi_{\max}}, \tag{8}$$

$$Z_\phi = k\frac{S'_{\max} - S'_\phi}{S'_{\max} + S'_{\min}}. \tag{9}$$

Note in this analysis that the depth (Z_ϕ) of each point on a cross-section varies from 0 to the value of scaling contrast times a constant k. The horizontal position of each point (X_ϕ) is linearly scaled with visual angle and is normalized by the maximum visual angle ($\phi_{\max}$) so that the range of positions varies between -1 and 1. An alternative procedure that produces fits only slightly worse than those obtained using Equation 8 is to scale X_ϕ with respect to position in the image plane (i.e., $X_\phi = X'_\phi / X'_{\max}$).

The dashed blue curves in Figure 13 show the estimated surface cross-sections based on Equations 8 and 9. It is important to keep in mind that there were only two free parameters in this analysis to maximize the goodness of fit among all of the different conditions. Equation 9 contains a depth scaling parameter (k), which was assigned a fixed value of 2.5 for the generation of each curve. The other free parameter is the restricted 45° range of visual angles that was considered for the analysis of convex surfaces (see Figure 9). Note in Figure 13 how closely the estimated shapes approximate the observers' settings in all of the different conditions. This model can account for the apparent depth-to-width ratios of the depicted surfaces, the apparent outward bowing of the surface asymptotic lines for convex surfaces, and the attenuated curvature at the near point for convex surfaces.

Discussion

There have been many previous studies reported in the literature that have attempted to assess which local attributes of texture (e.g., length, width, area, density, or foreshortening) are most informative for the perception of three-dimensional shape. One popular paradigm for addressing this issue is to place different texture cues in conflict with one another, either within static monocular images (Attneave & Olson, 1966; Braunstein & Payne, 1969; Cutting & Millard, 1984; Phillips, 1970; Rosenholtz & Malik, 1997; Todd & Akerstrom, 1987) or stereoscopic displays in which the texture information can also be in conflict with binocular disparity (Buckley, Frisby, & Blake, 1996; Cumming, Johnston, & Parker, 1993; Frisby & Buckley, 1992). The displays presented at inappropriate viewing angles in this study can be considered as a type of cue conflict paradigm because the information provided by scaling gradients and by foreshortening in those displays were inconsistent with one another. Another technique for investigating this issue is to manipulate the relative reliability of the cues by adding random variations to some local texture properties but not others (e.g., Knill, 1998a, 1998b). The variable size dot texture used in the present experiment (see Figure 7) is a good example of this type of manipulation. To the extent that they matter, the size variations should impact the analysis of texture scaling but have little or no effect on the analysis of foreshortening.

Following a spate of research on this topic in the late 1990s, many researchers concluded that texture foreshortening is the primary source of information for the perception of three-dimensional shape from texture, except perhaps in the special case of linear perspective where an image contains a pattern of converging contours (for a review, see Knill, 1998b). However, more recent research has provided a growing body of evidence that casts doubt on this hypothesis (Li & Zaidi, 2000, 2001; Saunders & Backus, 2006; Todd, Oomes, Koenderink, & Kappers, 2004; Todd et al., 2005; Zaidi & Li, 2002), and the results of the present experiment are particularly damaging to it.

If perceived three-dimensional shape were based on texture foreshortening, then reducing the visual angle with which an image is viewed should cause a corresponding reduction in perceived depth, but the actual empirical results are in the opposite direction (see Figure 9). It is important to point out that there have been several previous studies that have investigated the effects of viewing angle on perceived surface relief (Backus & Saunders, 2006; Farber & Rosinski, 1978; Lumsden, 1983; Nichols & Kennedy, 1993; Smith, 1967), although none of the previous discussions of this issue have considered the theoretical significance of that manipulation for the analysis of texture foreshortening. Although the size of the effect may vary, all of these studies agree that when the viewing angle for a fixed visual image is decreased, there is a corresponding increase in the magnitude of perceived depth.

In principle, when an image of a surface with isotropic texture is observed from the correct visual angle, an analysis of the optical foreshortening field is capable of producing a veridical estimate of the depicted three-dimensional shape. However, the empirical evidence on human perception demonstrates quite clearly that

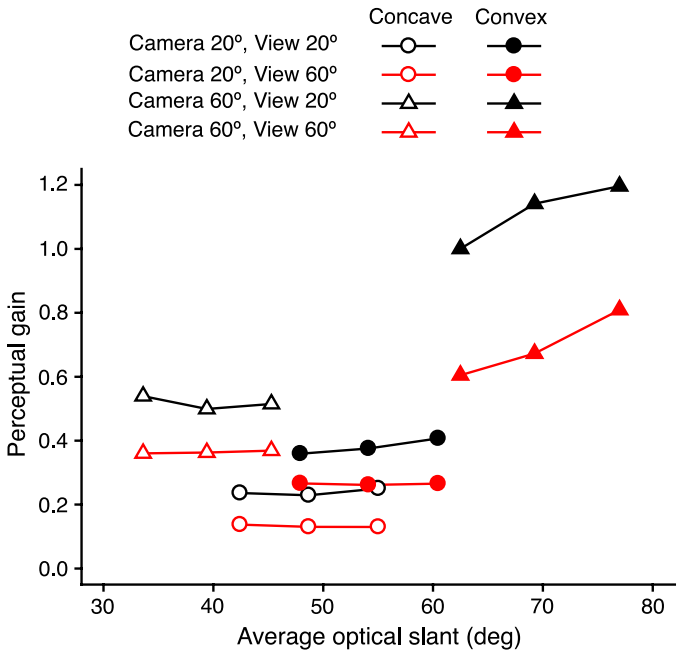

Figure 14. The perceptual gain of the observers' judgments as a function of the average optical slant for all possible combinations of viewing angle, camera angle, and sign of surface curvature.

observers' judgments of depth from texture are systematically underestimated (e.g., Backus & Saunders, 2006; Tibau et al., 2001; Todd & Akerstrom, 1987; Todd et al., 2004, 2005). This could potentially be explained by a perceptual bias or by the effects of conflicting flatness cues (e.g., see Backus & Saunders, 2006; Knill & Saunders, 2003), but neither of these influences can account for why convex surfaces appear to have greater relief than concave surfaces, or why the overall perceptual gain is influenced by the camera angle with which an image is rendered (see Figure 6).

A reviewer of an earlier draft of this manuscript suggested that these latter effects could perhaps be explained if the effects of biases or flatness cues systematically vary as a function of optical slant. Prior results have shown that differences among large surface slants are more reliably discriminated than are comparable differences among smaller slants (Knill, 1998a; Knill & Saunders, 2003). Thus, it might be reasonable to speculate that observers' reliance on biases or flatness cues may be diminished with increasing slant, as texture information becomes more and more reliable. If that were the case, then one might expect that lower slants may be more severely underestimated than higher slants. In an effort to test this hypothesis, we plotted the perceptual gain as a function of the average optical slant for all of the different combinations of viewing angle, camera angle, and sign of surface curvature (see Figure 14). Although there was a positive correlation between these variables, it accounts for only 40% of the variance. Note in particular that the effect of camera angle for the concave surfaces is in the opposite direction from what would be expected from a variable bias hypothesis. Another serious problem with this suggestion is the consequences it would have for the perception of planar surfaces. Because the optical slants on a planar surface vary in unison with visual angle in the direction of slant (Todd et al., 2005), a variable bias hypothesis would predict that planar surfaces viewed at large visual angles should appear highly curved.[3]

Another potential source of information for computing three-dimensional shape from texture is provided by local gradients of texture scaling as described by Equation 2 (Clerc & Mallat, 2002; Gårding, 1992, 1993; Malik & Rosenholtz, 1994, 1997; Purdy, 1958). A gradient-based analysis can account for the fact that reductions in viewing angle for a fixed visual image produce a systematic increase in the apparent depth-to-width ratio of a depicted surface, although it greatly overestimates the magnitude of this effect. For singly curved surfaces with homogeneous textures, like those used in the present experiment, a gradient-based analysis is capable of producing a veridical estimate of the depicted three-dimensional shape. As was the case for foreshortening, this type of analysis could be supplemented with a perceptual bias to account for the fact that observers' perceptions of depth from texture are systematically underestimated, but it cannot easily explain why convex surfaces appear to have greater relief than concave surfaces, or why the overall perceptual gain is influenced by the camera angle with which an image is rendered.

Although the results of the present experiment seem fundamentally incompatible with existing computational models for the perception of three-dimensional shape from texture, the overall pattern data can be fit quite

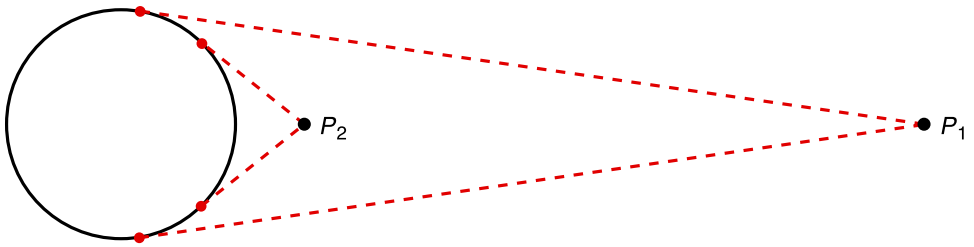

Figure 15. The visible portions of a sphere from two different vantage points P_1 and P_2. Note how the visible range of depths increases with viewing distance.

closely with a relatively simple analysis of scaling contrast. It should also be pointed out, however, that texture scaling is inversely related to the density of optical texture perpendicular to the direction of slant, and that the contrast of density is formally equivalent to the contrast of scaling. Thus, another reasonable hypothesis from these data is that the judged relative depths were determined from an analysis of texture density or spatial frequency in the appropriate direction (see also Thaler, Todd, & Dijkstra, 2007; Todd et al., 2005). These hypotheses can account for all of the different stimulus factors that influenced the observers' perceptions. They can accurately predict the relative magnitude of apparent depth over changes in the viewing angle, the camera angle, the sign of surface curvature, and the depicted surface depth. In addition, they can also predict the apparent outward bowing of the surface asymptotic lines for convex surfaces, and the attenuated curvature at the near point for convex surfaces.

Despite the impressive performance of the scaling contrast model for explaining the results of this study, there is an important limitation of the model that deserves to be highlighted. To exploit variations in texture scaling for determining apparent three-dimensional shape, the scaling variations must be large enough to be reliably detected. For surfaces that are planar or asymptotically planar, like the ones used in the present experiment, the variations in scaling can be arbitrarily large. That is not the case, however, for surfaces that are more continuously curved. To demonstrate this more clearly, it is useful to consider the optical projection of a sphere, as is schematically represented in Figure 15. As the viewing distance to a sphere becomes smaller and smaller, the increase in scaling contrast is limited because more and more of the surface becomes self-occluded.

If texture scaling is the primary source of information for the perceived three-dimensional shape of any particular surface configuration, then its apparent relief should be greatly reduced or eliminated when the surface is rendered under orthographic projection to remove all

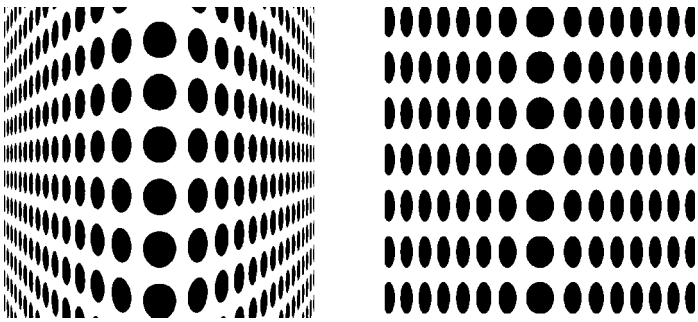

Figure 16. A hyperbolic cylinder under perspective projection with a 30° viewing angle (left) and under orthographic projection (right).

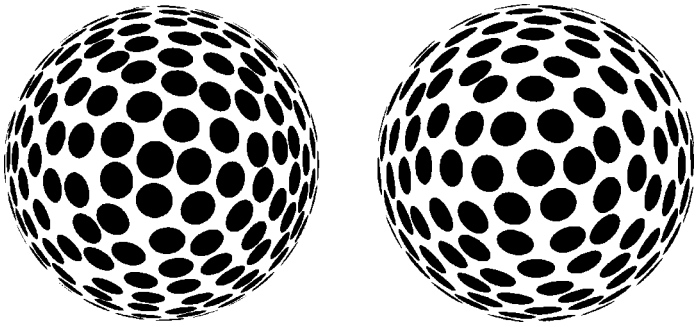

Figure 17. An ellipsoid surface under perspective projection with a 30° viewing angle (left) and under orthographic projection (right).

variations in scaling. That is exactly what occurs with images of planes or hyperbolic cylinders (see Figure 16). However, there are other surface configurations for which there is a compelling perception of shape from texture, even under orthographic projection (e.g., see Figure 17), and the apparent depths of those surfaces are relatively unaffected by the addition of perspective (see Todd & Akerstrom, 1987; Todd & Oomes, 2002). These observations suggest that texture scaling may be the primary source of information for surfaces that are close to planar or asymptotically planar viewed with large amounts of perspective, but that some other aspect of texture must also be used in other contexts.

Could that other aspect of texture be foreshortening? We suspect not, based on the findings of Todd et al. (2004) using anisotropic volumetric textures. Figure 18 shows two textured objects that are similar to the ones employed in that study. Note that the apparent near points on these surfaces all contain texture elements with elliptical shapes, and that the circular texture elements are all located in regions that appear slanted in depth. These perceptions would not be possible if the apparent

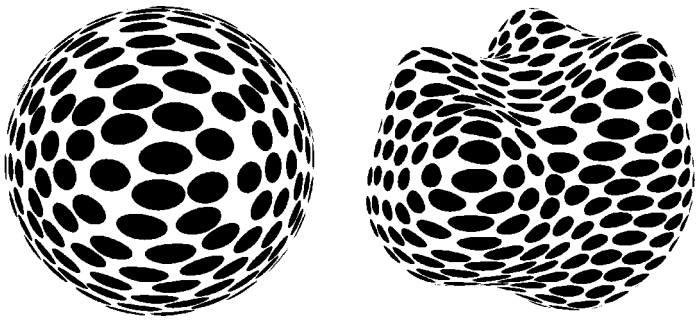

Figure 18. Two surfaces with anisotropic volumetric textures. These images were created using a volumetric texture that consisted of an array of small horizontally oriented ellipsoids whose centers were constrained to lie on the depicted surface. Any region of a surface that cut through an ellipsoid was colored black, and any region that cut through the space between ellipsoids was colored white.

local slants were determined by texture foreshortening. It should also be noted that the apparent depth extrema in any given direction are located in regions where the widths of the texture elements in that direction are local maxima (see also Grossberg & Mingolla, 1985, 1987; Todd & Akerstrom, 1987). We suspect this may be an important source of information for the perception of three-dimensional shape from texture with low levels of perspective, but the empirical support for that hypothesis will remain as an interesting problem for future research.

Acknowledgments

This research was supported by a grant from NSF (BCS-0546107).

Commercial relationships: none.
Corresponding author: James T. Todd.
Email: todd.44@osu.edu.
Address: Department of Psychology, The Ohio State University, 1827 Neil Avenue, Columbus, OH 43210, USA.

Footnotes

[1]In an extensive series of experiments, Li & Zaidi, (2000, 2001) have consistently shown that observers are unable to accurately perceive the sign of curvature for surfaces with isotropic textures. Because several of the textures in the present experiment were statistically isotropic, the accuracy of observers' judgments may appear at first blush to be fundamentally incompatible with those earlier studies. The solution to this conundrum is suggested by Todd et al. (2005) who showed that observers can accurately determine the sign of curvature for surfaces with isotropic texture if the visual angle of the depicted surface patch is sufficiently large (see also Saunders & Backus, 2006). The camera angles used in this study were all well above threshold for performing this type of judgment.

[2]The values of scaling contrast shown in Figure 12 were computed from the projected lengths of idealized texture elements along a horizontal cross-section through the center of each image and did not take into account the variability of the physical texture elements in the variable size dot and the variable shape ellipse textures. To obtain a stable measure of texture scaling when there are random variations among the physical texture elements, it is necessary to average the projected sizes of the optical texture elements over an appropriately large neighborhood. A more detailed discussion of this issue is presented in Thaler et al. (2007).

[3]Consider, for example, a planar surface at a 50° slant relative to the frontoparallel plane that is observed with a viewing angle of 60° (see Figure 1). The local optical slants across different regions of the surface in that case would vary from 20° to 80°. To account for the results of this study using a variable bias hypothesis, the systematic underestimation of slant would have to vary between 20% and 80% over that range of optical slants (see Figure 14). The change in apparent curvature that would cause is much larger than what is evident in observers' judgments.

References

Aloimonos, J. (1988). Shape from texture. *Biological Cybernetics, 58,* 345–360. [PubMed]

Attneave, F., & Olson, R. K. (1966). Inferences about Visual Mechanisms from Monocular Depth Effects. *Psychonomic Science, 4,* 133–134.

Backus, B. T., & Saunders, J. A. (2006). The accuracy and reliability of perceived depth from linear perspective as a function of image size. *Journal of Vision, 6*(9):7, 933–954, http://journalofvision.org/6/9/7/, doi:10.1167/6.9.7. [PubMed] [Article]

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

Blake, A., & Marinos, C. (1990). Shape from texture: Estimation, isotropy and moments. *Artificial Intelligence, 45,* 323–380.

Blostein, D., & Ahuja, N. (1989). Shape from texture: Integrating texture-element extraction and surface estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 11,* 1233–1251.

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

Brown, L. G., & Shvayster, H. (1990). Surface orientation from projective foreshortening of isotropic texture autocorrelation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 12,* 584–588.

Buckley, D., Frisby, J. P., & Blake, A. (1996). Does the human visual system implement an ideal observer theory of slant from texture? *Vision Research, 36,* 1163–1176. [PubMed]

Clerc, M., & Mallat, S. (2002). The texture gradient equation for recovering shape from texture. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 24,* 536–549.

Cumming, B. G., Johnston, E. B., & Parker, A. J. (1993). Effects of different texture cues on curved surfaces

viewed stereoscopically. *Vision Research, 33,* 827–838. [PubMed]

Cutting, J. E., & Millard, R. T. (1984). Three gradients and the perception of flat and curved surfaces. *Journal of Experimental Psychology: General, 2,* 198–216. [PubMed]

Farber, J., & Rosinski, R. R. (1978). Geometric transformations of pictured space. *Perception, 7,* 269–282. [PubMed]

Frisby, J. P., & Buckley, D. (1992). Experiments on stereo and texture cue combination in human vision using quasi-neutral viewing. In G. A. Orban, & H. H. Nagel (Eds.), *Artificial and biological vision systems* (pp. 268–297). Berlin: Springer.

Gårding, J. (1992). Shape from texture for smooth curved surfaces in perspective projection. *Journal of Mathematical Imaging and Vision, 2,* 327–350.

Gårding, J. (1993). Shape from texture and contour by weak isotropy. *Artificial Intelligence, 64,* 243–297.

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

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

Gibson, J. J. (1987). The perception of visual surfaces. *American Journal of Psychology, 100,* 646–664. [PubMed]

Grossberg, S., & Mingolla, E. (1985). Neural dynamics of perceptual grouping: Textures, boundaries and emergent segmentations. *Perception & Psychophysics, 38,* 141–171. [PubMed]

Grossberg, S., & Mingolla, E. (1987). Neural dynamics of surface perception: Boundary Webs, illuminants, and shape from shading. *Computer Vision, Graphics, and Image Processing, 37,* 116–165.

Knill, D. C. (1998a). Discrimination of planar surface slant from texture: Human and ideal observers compared. *Vision Research, 38,* 1683–1711. [PubMed]

Knill, D. C. (1998b). Surface Orientation from texture: Ideal observers, generic observers and the information content of texture cues. *Vision Research, 38,* 1655–1682. [PubMed]

Knill, D. C., & Saunders, J. A. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research, 43,* 2539–2588. [PubMed]

Krumm, J., & Shafer, S. A. (1992). Shape from periodic texture using the spectrogram. *Proceedings of the 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '92),* 284–289.

Li, A., & Zaidi, Q. (2000). Perception of three-dimensional shape from texture is based on patterns of oriented energy. *Vision Research, 40,* 217–242. [PubMed]

Li, A., & Zaidi, Q. (2001). Verdicality of three-dimensional shape perception predicted from amplitude spectra of natural textures. *Journal of the Optical Society of America A, Optics, image science, and vision, 18,* 2430–2447. [PubMed]

Lumsden, E. A. (1983). Perception of radial distance as a function of magnification and truncation of depicted spatial layout. *Perception & Psychophysics, 33,* 177–182. [PubMed]

Malik, J., & Rosenholtz, R. (1994). A computational model for shape from texture. *Ciba Foundation Symposium, 84,* 272–283. [PubMed]

Malik, J., & Rosenholtz, R. (1997). Computing local surface orientation and shape from texture for curved surfaces. *International Journal of Computer Vision, 23,* 149–168.

Marinos, C., & Blake, A. (1990). Shape from texture: The homogeneity hypothesis. *Proceedings of the 3rd International Conference on Computer Vision* (pp. 350–334). Osaka, Japan.

Nichols, A. L., & Kennedy, J. M. (1993). Angular subtense effects on perception of polar and parallel projections of cubes. *Perception & Psychophysics, 54,* 763–772. [PubMed]

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

Purdy, W. P. (1958). The hypothesis of psychophysical correspondence in space perception. *Dissertation Abstracts International, 42,* 1454. (University Microfilms No. 58–5594).

Rosenholtz, R., & Malik, J. (1997). Surface orientation from texture: Isotropy or homogeneity (or both)? *Vision Research, 37,* 2283–2293. [PubMed]

Sakai, K., & Finkel, L. H. (1995). Characterization of spatial frequency spectrum in the perception of shape from texture. *Journal of the Optical Society of America A, Optics, image science, and vision, 12,* 1208–1224. [PubMed]

Saunders, J. A., & Backus, B. T. (2006). Perception of surface slant from oriented textures. *Journal of Vision, 6*(9):3, 882–897, http://journalofvision.org/6/9/3/, doi:10.1167/6.9.3. [PubMed] [Article]

Smith, A. H. (1967). Perceived slant as a function of stimulus contour and vertical dimension. *Perceptual and Motor Skills, 24,* 167–173.

Super, B. J., & Bovik, A. C. (1995). Planar surface orientation from texture spatial frequencies. *Pattern Recognition, 28,* 729–743.

Thaler, L., Todd, J. T., & Dijkstra, T. M. (2007). The effects of phase on the perception of 3D shape from texture: Psychophysics and modeling. *Vision Research, 47,* 411–427. [PubMed]

Tibau, S., Willems, B., Van Den Berg, E., & Wagemans, J. (2001). The role of the center of projection in the estimation of slant from texture of planar surfaces. *Perception, 30,* 185–193. [PubMed]

Todd, J. T., & Akerstrom, R. A. (1987). Perception of three dimensional form from patterns of optical texture. *Journal of Experimental Psychology: Human Perception and Performance, 2,* 242–255. [PubMed]

Todd, J. T., & Oomes, A. H. (2002). Generic and non-generic conditions for the perception of surface shape from texture. *Vision Research, 42,* 837–850. [PubMed]

Todd, J. T., Oomes, A. H., Koenderink, J. J., & Kappers, A. M. (2004). The Perception of doubly curved surfaces from anisotropic textures. *Psychological Science, 15,* 40–46. [PubMed]

Todd, J. T., Thaler, L., & Dijkstra, T. M. (2005). The effects of field of view on the perception of 3D slant from texture. *Vision Research, 45,* 1501–1517. [PubMed]

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

Zaidi, Q., & Li, A. (2002). Limitations on shape information provided by texture cues. *Vision Research, 42,* 815–835. [PubMed]