The perception of 3D shape from texture based on directional width gradients

James T. Todd

Department of Psychology, Ohio State University,
Columbus, OH, USA

Lore Thaler

Department of Psychology, University of Western Ontario,
London, ON, Canada

A new computational analysis is described that is capable of estimating the 3D shapes of continuously curved surfaces with anisotropic textures that are viewed with negligible perspective. This analysis assumes that the surface texture is homogeneous, and it makes specific predictions about how the apparent shape of a surface should be distorted in cases where that assumption is violated. Two psychophysical experiments are reported in an effort to test those predictions, and the results confirm that observers' ordinal shape judgments are consistent with what would be expected based on the model. The limitations of this analysis are also considered, and a complimentary model is discussed that is only appropriate for surfaces viewed with large amounts of perspective.

Keywords: shape, depth, slant, texture

Introduction

In his influential book, *The perception of the visual world*, Gibson (1950) introduced the concept of texture gradients as a potential source of optical information about the layout of surfaces in the environment. Since then, many different properties of optical texture have been identified that can be used to estimate various aspects of 3D surface structure (Gårding, 1992, 1993; Purdy, 1958), and many psychophysical experiments have been performed in an effort to determine which of these potential sources of information are most relevant to human perception. The literature on this topic contains many conflicting results, however, and there is no clear consensus as yet on how patterns of optical texture are perceptually analyzed by human observers.

One of the most popular computational models for estimating 3D shape from texture is based on an assumption that variations in reflectance on a surface are statistically isotropic—i.e., that they are approximately equal in all directions—and that anisotropic patterns of texture within a visual image can be attributed to variations in surface slant. These distortions are often referred to as texture foreshortening. They can be measured from the aspect ratios of individual optical texture elements (Stevens, 1981a, 1984), 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 (Bajcsy & Lieberman, 1976; Brown & Shvayster, 1990; Krumm & Shafer, 1992; Sakai & Finkel, 1995; Super & Bovik, 1995).

Although foreshortening is probably the most popular source of information in computational analyses of shape from texture, there is a growing body of evidence to indicate that it has little or no influence on observers' perceptions of surface slant (Backus & Saunders, 2006; Gillam, 1970; Todd, Christensen, & Guckes, 2010; Todd, Oomes, Koenderink, & Kappers, 2004; Todd, Thaler, & Dijkstra, 2005; Todd, Thaler, Dijkstra, Koenderink, & Kappers, 2007). Consider, for example, the image of a slanted plane under orthographic projection in the top panel of Figure 1. Any of the computational models cited above would be able to correctly determine that the depicted surface has a 50° slant relative to the fronto-parallel plane. However, when the same image is presented to human observers, they are unable to perceive any slant at all.

An alternative source of information for estimating shape from texture is based on an assumption that distributions of reflectance on a surface are statistically homogeneous—i.e., that they are the same at all surface locations—and that systematic gradients of optical texture across different regions of an image can therefore be attributed to variations in surface depth, slant or curvature. The first computational analysis of texture gradients was performed by Purdy (1958), who examined the spatial variations of several different aspects of optical texture elements, including their major axes, minor axes, areas, aspect ratios and density. He showed that the any of these
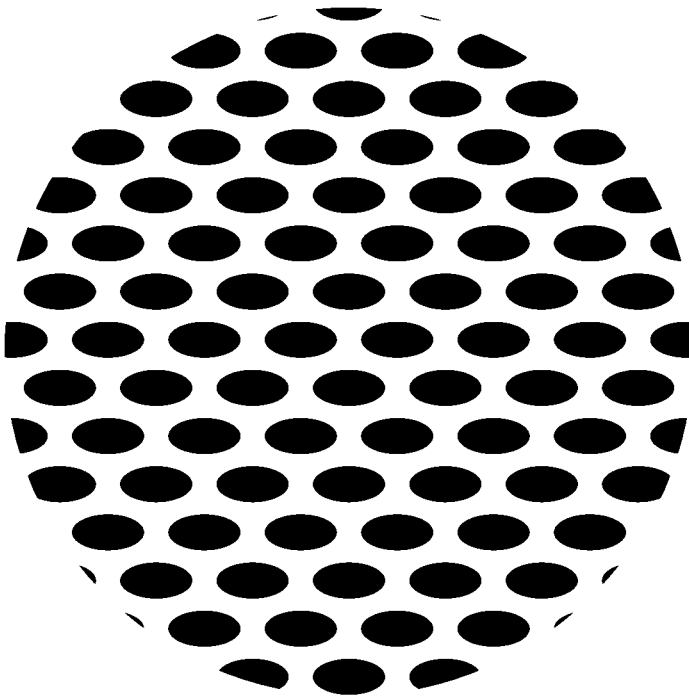
Figure 1. An image of a planar surface with a 50° slant that was rendered under orthographic projection.

properties $\mu(\phi, \gamma)$ at a point on a planar surface in a visual direction (ϕ, γ) can be used to determine the optical slant $\sigma(\phi, \gamma)$ at that point from the normalized gradient defined by the following equation:

$$\tan \sigma(\phi, \gamma) = k \frac{\nabla \mu(\phi, \gamma)}{\mu(\phi, \gamma)}, \qquad (1)$$

where k is a constant that is peculiar to each type of gradient. Gårding (1992) later extended this analysis to include curved surfaces. He showed that the normalized major axis gradient is the only one that that can be used to estimate local surface slant independently of surface curvature. This is often referred to as a scaling gradient. It can be measured by calculating the relative lengths of texture elements in different regions of an image (Purdy, 1958), or from the affine correlations between the amplitude spectra in those regions (Clerc & Mallat, 2002; Malik & Rosenholtz, 1994, 1997).

A recent series of studies by Todd et al. (2005, 2007, 2010) has shown that the ability to distinguish slants from texture for planar or asymptotically planar surfaces requires relatively large viewing angles, which provides strong evidence that observers' perceptions cannot be based on computational analyses within small local neighborhoods, such as spatial derivatives. In light of this finding, they proposed that perceived shape from texture for this class of surfaces is based on an alternative source

of information called scaling contrast, as defined by the following equation:

$$Z(\phi, \gamma) \propto \frac{\lambda_{\text{Max}} - \lambda(\phi, \gamma)}{\lambda_{\text{Max}} + \lambda_{\text{Min}}}, \qquad (2)$$

where $Z(\phi, \gamma)$ is the relative depth of a surface point in a visual direction (ϕ, γ), $\lambda(\phi, \gamma)$ is the projected major axis of a texture element at that point, and λ_{Max} and λ_{Min} are the maximum and minimum major axis lengths. Note that this is similar to a normalized gradient, but it is designed to evaluate the variations in scaling over large regions of visual space, rather than small local neighborhoods.

The scaling contrast model predicts that slanted surfaces presented under orthographic projection should have no apparent depth at all because there are no variations in texture scaling. For slanted surfaces under perspective projection, the model predicts that the magnitude of perceived relief should increase systematically with the depicted field of view as has been demonstrated by Backus and Saunders (2006) and Todd et al. (2005, 2007, 2010). It can also account for differences in perceived relief between concave and convex surfaces (Todd et al., 2005, 2007), and the perceptual distortions that occur when images of textured surfaces are observed from an inappropriate viewing distance (Backus & Saunders, 2006; Todd et al., 2007, 2010).

If scaling contrast is the primary source of information for the perception of 3D shape from texture, then the magnitude of perceived relief should be reduced to zero when a surface is rendered under orthographic projection so as to remove all variations in scaling. That is exactly what occurs with images of slanted planes as is shown in Figure 1. However, continuously curved surfaces can produce a compelling perception of shape from texture, even under orthographic projection (see Figure 2), 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 planar or asymptotically planar, but that some other aspect of texture must also be used in other contexts.

Could that other aspect of texture be foreshortening? The perceptual appearance of the surface in Figure 2 provides strong evidence that this is probably not the case. This surface has an anisotropic texture that is elongated in a vertical direction. If any of the foreshortening models cited above were applied to this image, all of the points along a vertical cross-section through its center would have an estimated slant of 41°, and they would all be interpreted as inflexion points in the depth profiles of the horizontal cross-sections. That is not how the surface is perceived, however. Despite the fact that the texture elements in the central axis of the cylinder are all
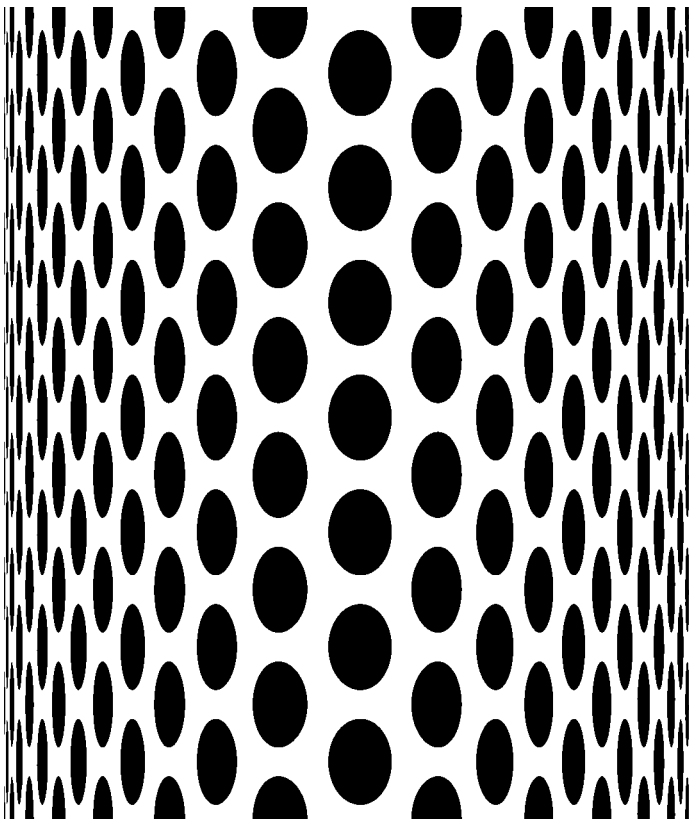
Figure 2. An image of an elliptical cylinder with an anisotropic texture that was rendered under orthographic projection.

noticeably foreshortened, they appear nonetheless to have a frontoparallel orientation, and they are perceptually interpreted as local depth minima rather than inflection points.

Todd and Akerstrom (1987) investigated observers' perceptions for displays that were quite similar to Figure 2, and concluded that the primary source of information must be the systematic variations of texture element widths in the direction of slant. Fleming, Torralba, and Adelson (2004) have more recently pointed out that the pattern of texture compression along an arbitrary surface cross-section has a functional form that is quite similar to the pattern of surface slant. In the discussion that follows we will extend these observations to develop a specific computational model for estimating the 3D shapes of objects like the one in Figure 2.

The basic insight underlying this model is that the optical projections of texture elements at local depth extrema along a surface cross-section can be used as anchors to estimate the local slants of all other points that are aligned in the same direction. Figure 3 shows a narrow cross-section of the image in Figure 2, together with the depth profile of the depicted surface. The projected width of a texture element at a position ϕ on the cross-section is labeled $\omega(\phi)$, and the local width maximum that occurs at the depth extremum is labeled ω_{Max}. Consider a vector

$\mathbf{N}(\phi)$ in the plane of the surface cross-section that is normal to its outer boundary at the position ϕ. If the surface is depicted with negligible perspective and its texture is homogeneous, then the angle $\sigma(\phi)$ of this vector relative to the z-axis can be determined by the following equation:

$$\cos \sigma(\phi) = \frac{\omega(\phi)}{\omega_{\text{Max}}}. \qquad (3)$$

It is important to recognize that the vector $\mathbf{N}(\phi)$ will not in general be a surface normal. Rather, it is the component of the surface normal that is in the plane of the cross-section. Similarly, surface points at the local depth extrema will not generically have fronto-parallel orientations. They are the points along a cross-section that are closer to fronto-parallel than their neighbors. It should also be noted that edge density along a cross-section is the reciprocal of texture width, and could therefore provide an alternative measure for calculating slant in Equation 3.
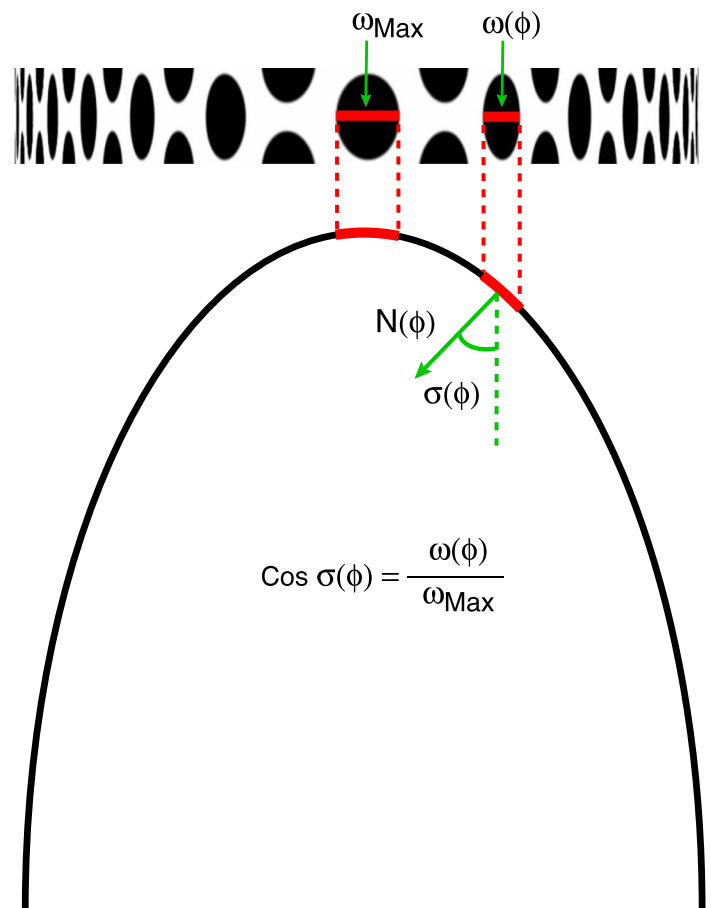

Figure 3. A horizontal cross-section of the texture pattern in Figure 2 and the depth profile of the depicted surface. The slant of the vector $\mathbf{N}(\phi)$ relative to the z-axis can be determined by the projected texture width at position ϕ relative to the projected texture width at the local depth minimum.

Indeed, this would be the preferred approach for surfaces with more fine scale textures whose elements are not easily individuated (see Thaler, Todd, & Dijkstra, 2007).

In order to effectively implement this model using either widths or densities it is necessary to identify the local depth extrema along arbitrary surface cross-sections. Todd et al. (2004) have shown that observers can do this quite accurately, even in some cases when the depicted surface has an inhomogeneous texture. The stimuli in that study were created using a technique called volumetric texturing, which is analogous to sculpting an object out of a solid material, such as wood or marble. A schematic illustration of this process is provided in Figure 4A. This texture was created using a set of spheres that were distributed in 3D space such that their centers were constrained to lie on an ellipsoid surface. Any region of the ellipsoid that cut through a sphere was colored black, and any region that cut through the space between spheres was colored white. This produces a pattern of circular polka dots on the surface that is statistically homogeneous and isotropic. Local depth minima along surface cross-sections can be identified in this case by local maxima in the widths of the texture elements in the direction of the cross-section, or by local minima of texture foreshortening.

The images presented in Figures 4B and 4C show the same ellipsoid surface with volume textures that were created using a set of ellipsoids rather than spheres. For the surface shown in Figure 4B, the texture ellipsoids were oriented in a horizontal direction. Note in this case that the overall pattern of perceived relief is completely incompatible with what would be expected from an analysis of foreshortening gradients, because the regions that appear closest in depth are not the ones whose optical texture elements are closest to circular. The perceptual appearance is consistent, however, with what would be expected from an analysis of directional width gradients. That is to say, along any given surface cross-section in any given direction, the apparent near point is always located in a region where the width of the texture elements in that direction is a local maximum. Even though this type of texture is inhomogeneous, an analysis of directional width gradients will produce qualitatively accurate estimates of 3D shape whenever the ellipsoids that define the texture elements are oriented either parallel or
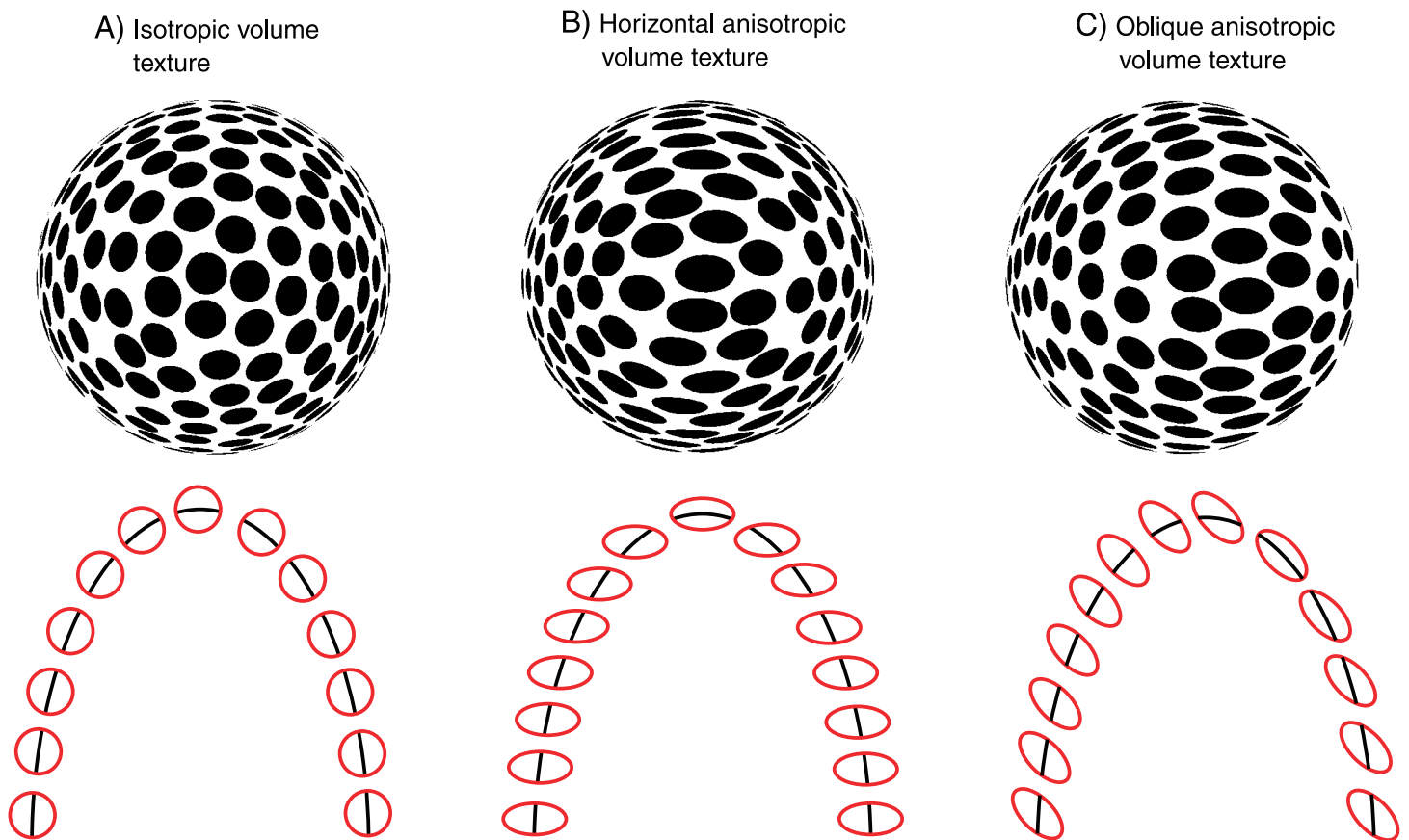

Figure 4. Images of an ellipsoid surface with three different types of volume textures. A) An isotropic texture composed of small spheres that are centered on the depicted surface; B) an anisotropic texture composed of small ellipsoids oriented perpendicular to the line of sight; and C) an anisotropic texture composed of small ellipsoids oriented at an oblique angle to the line of sight.

perpendicular to the surface cross-section an observer is asked to judge. These estimates will be systematically distorted, however, when the ellipsoids used to generate the texture are oriented in depth at an oblique angle to the designated cross-section. An example of this condition is shown in Figure 4C. Note in this case that the object appears perceptually to be slanted to the right, in precisely the same way that would be theoretically predicted by an analysis of its directional width gradients. The research described in the present article was designed to provide more rigorous empirical support for these anecdotal observations.

Experiment 1

Methods

Subjects

Seven observers participated in the experiment, including the authors, and five others who were naïve about the issues being investigated. All of the observers had normal or corrected-to-normal visual acuity.

Apparatus

The experiment was conducted using a Dell Dimension 8300 PC with an ATI Radeon 9700 PRO graphics card, and a 19 in CRT with a spatial resolution of 1280×1024 pixels. The stimulus images were presented within a 29.3 cm^2 region of the display screen, which subtended 20° of visual angle when viewed at a distance of 83 cm. The displays were viewed monocularly with an eye patch, and a chin rest was used to constrain head movements.

Stimuli

The stimuli depicted two possible ellipsoid surfaces, both of which had circular occlusion boundaries that subtended 15° of visual angle (see Figure 5). The visible portion of one of these objects had aspect ratios of 1.37 (depth) $\times$ 1 (width) $\times$ 1 (height), and its major axis was coincident with the line of sight. To create the second object, we compressed the first one by 16% in a horizontal direction and rotated it 15° about a vertical axis. This transformed object had the same circular occlusion boundary as the first, but its local depth minimum was shifted horizontally such that its optical projection was located at an eccentricity of 3.75°, which was exactly halfway between the center of the object's projection and the occlusion boundary. These surfaces were textured using the same process described in Figure 4. The texture was composed of a set of small ellipsoids that were distributed in 3D space such that their centers were constrained to lie on the surface of a depicted object. Any region of the surface that cut through one of the small ellipsoids was colored black, and any region that cut through the space between ellipsoids was colored white. This produces a pattern of elliptical markings on the surface whose shapes vary systematically as a function of surface orientation. Note that by orientating the texture ellipsoids at an oblique angle in depth, it is possible to manipulate the horizontal position of the local width maximum independently of the position of the local depth minimum. Thus, there were four main experimental conditions with centered or shifted depth minima, and centered or shifted width maxima in all possible combinations. We created eight different distributions of texture for each of these conditions, half of which had shifts to the left as shown in Figure 5, and the other half had shifts to the right.
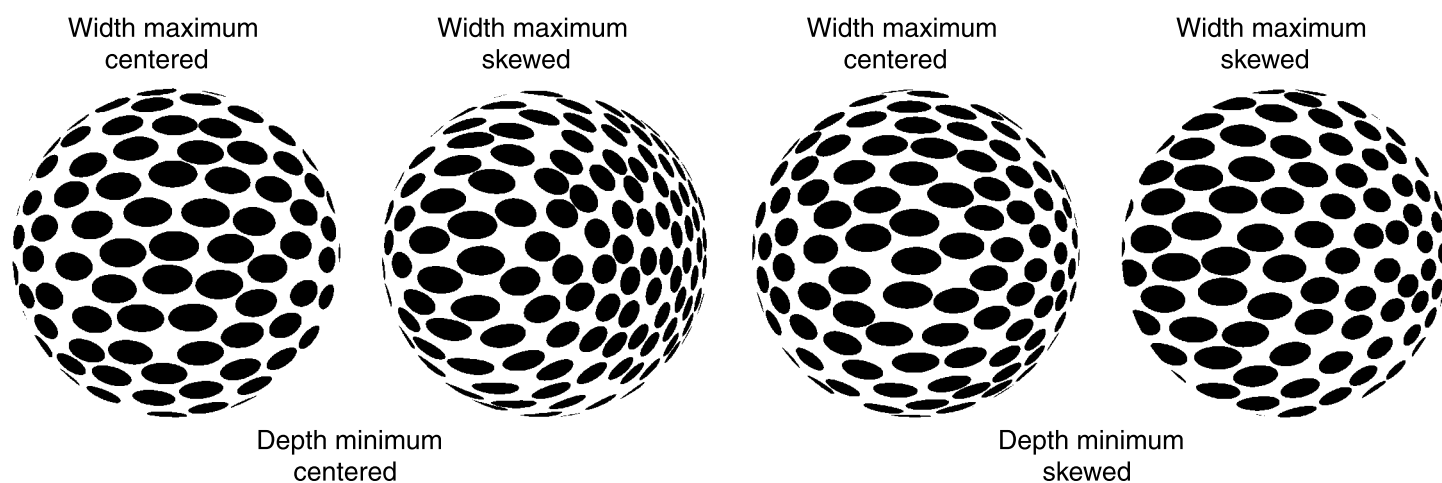

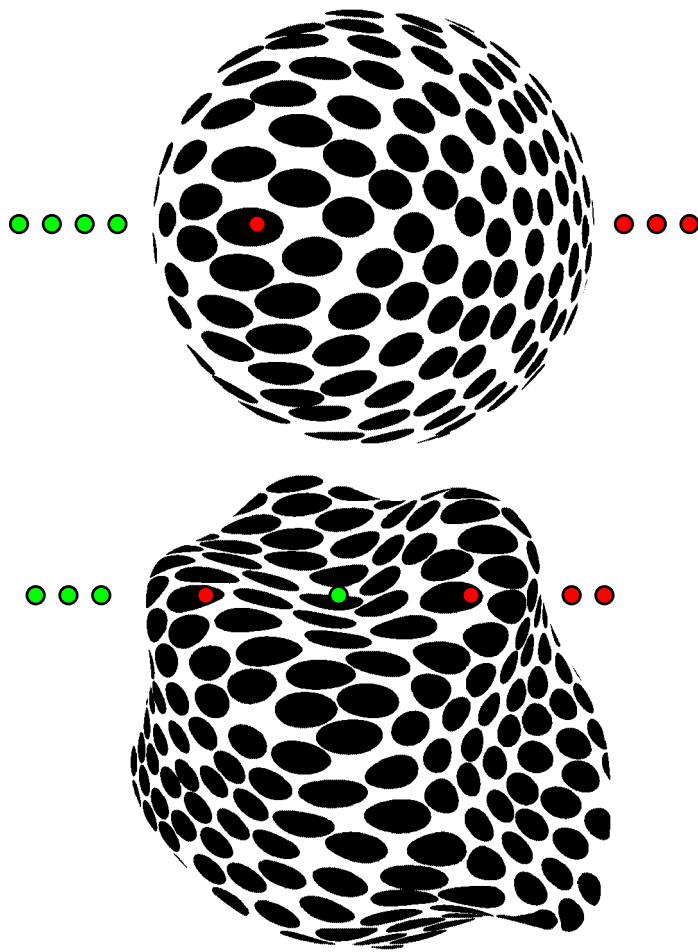Figure 5. The four stimulus conditions from Experiment 1.

Figure 6. The task for Experiments 1 and 2. Top Panel: In Experiment 1 observers adjusted the horizontal positions of small dots to mark the local depth extrema along a scan line through the center of an ellipsoid surface. Bottom Panel: In Experiment 2 observers marked the local depth extrema on more complex surface structures. The horizontal scan line on each trial could be located at one of four possible vertical positions within the image.

Procedure

On each trial an image of a textured surface was presented together with four green dots to the left and four red dots to the right that could be moved along a single horizontal scan line through the center of the image with a hand held mouse (see the top panel of Figure 6). Observers were instructed to mark each local depth minimum on the scan line with a red dot and each local depth maximum with a green dot, excluding the occlusion contours. Once all of the depth extrema on that scan line were appropriately marked, the trial was terminated by pressing a mouse button, and a new display was presented. Within each block of trials, all 32 of the possible stimulus images were presented in a random sequence. Each observer participated in 4 blocks within a single experimental session.

Results

Although they were free to mark multiple depth extrema along the surface scan lines, all of the observers marked just a single depth minimum on every trial, which indicates that all of the depicted surfaces were perceived as uniformly convex. Figure 7 shows the judged positions of these local depth minima averaged over all observers for each of the 32 possible stimuli. The top graph in this figure shows the average settings plotted as a function of the actual depth extrema, whereas the bottom graph presents the data as a function of the local width maxima. It is clear from these results that virtually all of the variance in observers' judgments among the different
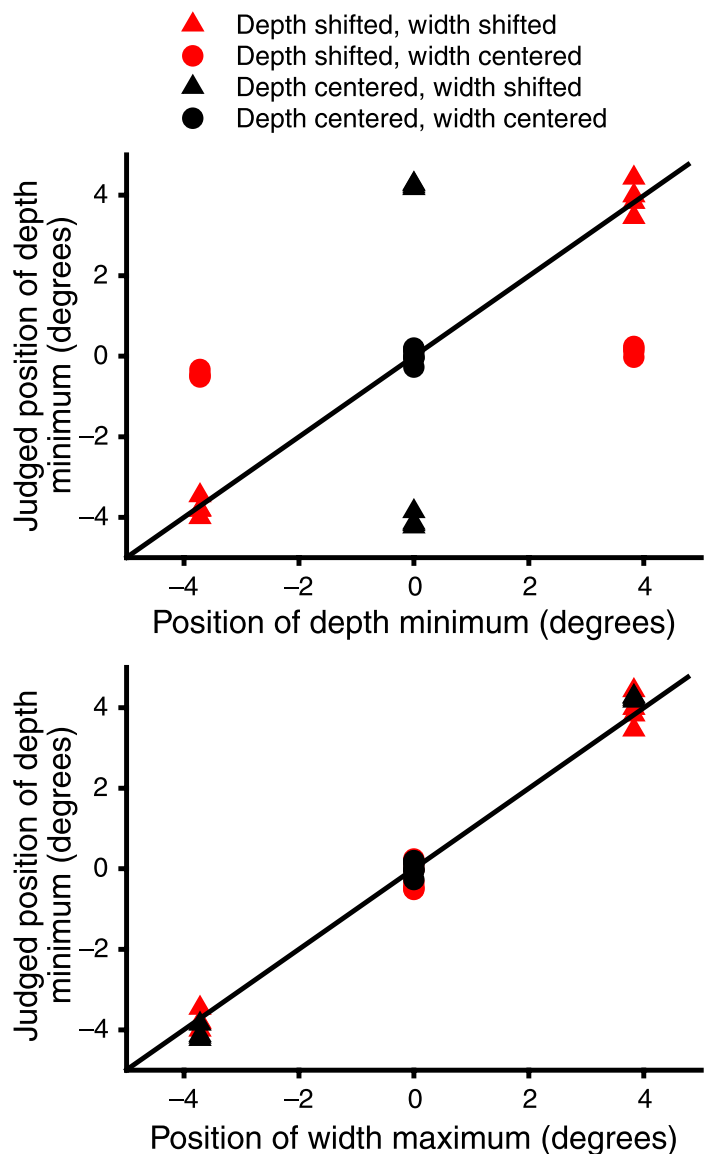

Figure 7. The average settings over all observers in Experiment 1 as a function of the ground truth (top) and the local width maxima (bottom).

possible stimuli can be accounted for by the positions of the local width maxima. Experiment 2 was designed to confirm this finding using textured images of more complex surface structures that contain both concavities and convexities.

Experiment 2

Methods

The apparatus and procedure were identical to those described for Experiment 1. The stimuli were images of three randomly deformed spheres (see Figure 8) that were textured using the same procedure as described in Figure 4. The texture was composed of a set of small ellipsoids that were distributed in 3D space such that their centers were constrained to lie on the surface of a depicted object. In the horizontal condition the ellipsoids were all aligned parallel to the image plane. In the horizontal-oblique condition they were rotated 45° about a vertical axis. We also presented these same images rotated 90° relative to the line of sight such that the depicted texture elements were all oriented vertically. These will be referred to as the vertical and vertical-oblique conditions. Four different distributions of texture were created for each combination of object and condition resulting in a total of 48 possible stimulus displays. These were presented to five of the observers from Experiment 1, including the two authors. On each trial an image of a textured surface was presented together with four green dots to the left and four red dots to the right that could be moved along a single horizontal scan line with a hand held mouse (see bottom panel of Figure 6). Observers were instructed to mark each local depth minimum on the scan line with a red dot and each local depth maximum with a green dot. Within a single experimental session, observers were asked to identify the local depth extrema along four different horizontal scan lines for each of the 48 possible stimuli.

Results

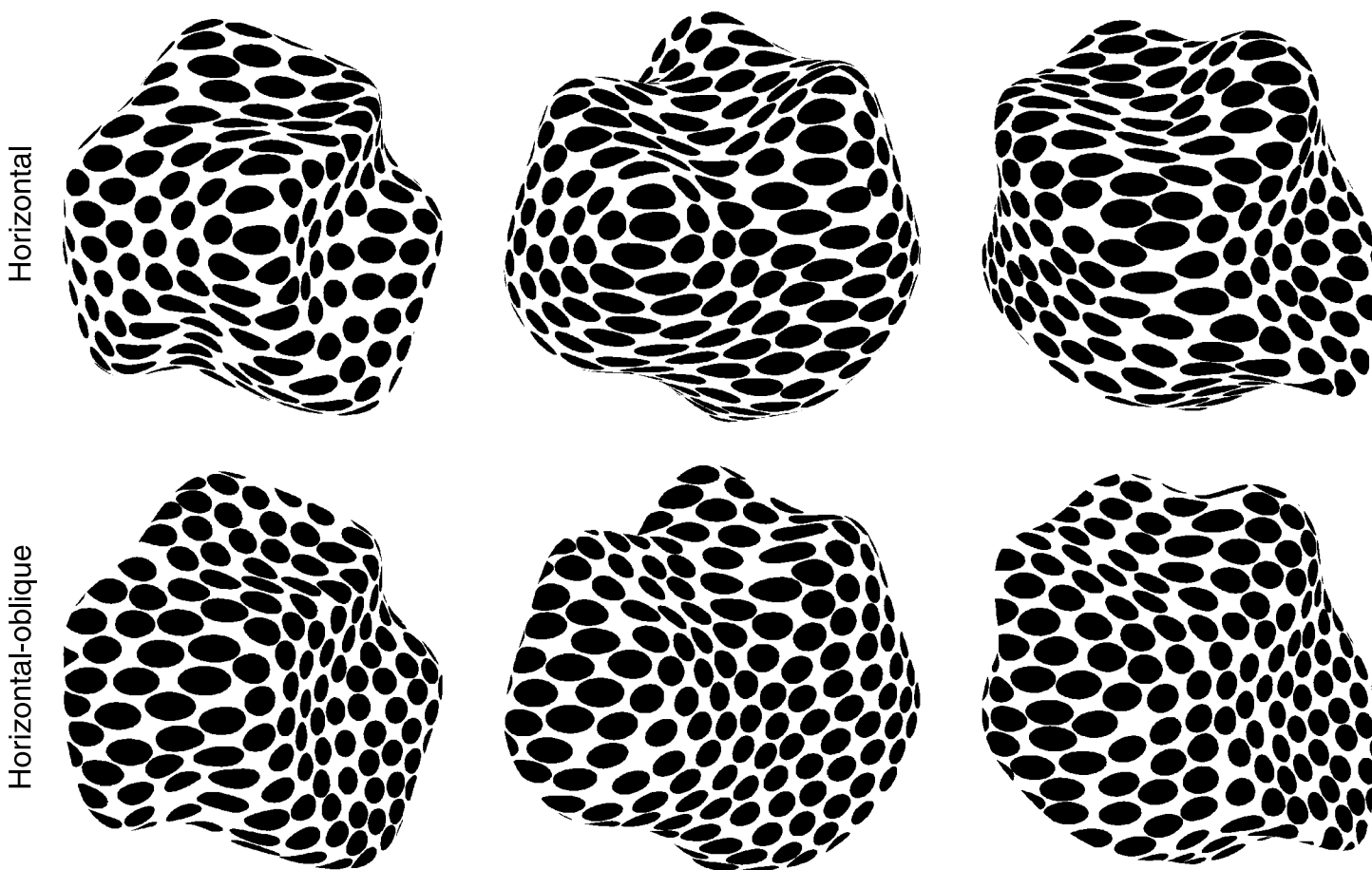To determine the width functions for these stimuli, we created a new set of images in which the ellipsoids were

Figure 8. The three possible objects from Experiment 2 with horizontal and horizontal-oblique volume textures. These same images were also rotated 90° about the line of sight in the vertical and vertical-oblique conditions.

all positioned so that their centers would project to points along the different possible scan lines for each combination of object and condition. This provided a dense sample of texture element widths along each scan line, which were fit using a cubic spline interpolation in SigmaPlot 11 to generate a continuous function. The red curves in Figure 9 shows the width functions obtained for two scan lines on one of the possible stimulus objects. The solid red curves in this figure show the width functions for the horizontal condition; the dashed red curves show the corresponding functions for the horizontal-oblique condition; and the black curves show the surface depth profiles. It is important to note that in the horizontal condition the local width maxima are all perfectly aligned with the local depth extrema and inflection points on the surface depth profile. This same pattern of alignment also occurs in the vertical and vertical-oblique conditions. That is not the case, however, in the horizontal-oblique condition. Note in Figure 9 how the horizontal-oblique width maxima can be displaced relative to the depth extrema, and that the magnitude of this displacement is greatest in surface regions that have the largest changes in depth.

The dots on each width function in Figure 9 show the average positions of the judged depth minima (solid dots) and maxima (open dots) over all of the different observers and texture distributions; the numbers just above them show the percentage of possible trials that each point was marked; and the horizontal bars depict the standard deviation of these judgments. It is interesting to note that observers often misinterpreted the inflection points as local depth minima, which resulted in false alarm responses. A typical example of this is evident in the lower scan line for both of the images in Figure 9. Observers also sometimes failed to detect shallow concavities on a surface, resulting in misses. These misses or false alarms occurred on approximately 13% of the trials, and both types of error always occurred in pairs involving one depth minimum and one depth maximum (see also Todd et al., 2004).

The top panel of Figure 10 shows the averaged judged positions of the depth extrema in all conditions (excluding false alarms) plotted as a function of the actual depth extrema. All judgments from the horizontal, vertical and vertical-oblique conditions are represented by black dots. An analysis of linear regression revealed that the ground truth accounts for 96% of the variance among the judgments in these conditions. The judgments obtained for the horizontal-oblique condition are highlighted in red, but their correlation with the ground truth is much lower, accounting for only 82% of the variance. The lower panel of Figure 10 shows the same data plotted as a function of the local width maxima. It is important to keep in mind that the local depth extrema are always coincident with local width maxima in the horizontal, vertical and vertical-oblique conditions, so that the correlation of observers' judgments with the local width maxima in those conditions is identical to their correlation with the ground truth. The two predictors diverge, however, for the judgments obtained in the horizontal-oblique condition, and the local width maxima in that case account for 97% of the variance. When considered in combination with the results of Experiment 1, these results provide compelling evidence that the pattern of texture width changes along
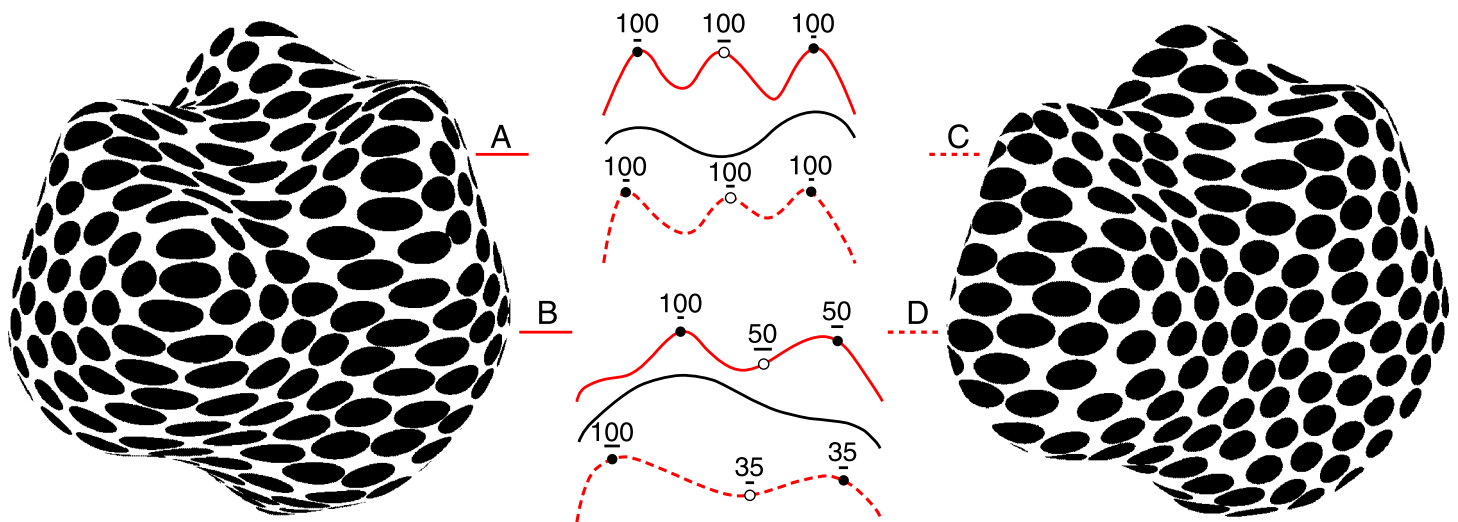

Figure 9. One of the possible stimulus objects from Experiment 2 with horizontal and horizontal-oblique textures. The solid and dashed red curves in the center panel show the width functions along two designated scan lines, and the black curves depict the surface depth profile along those scan lines. The dots on each width function show the average positions of the judged depth minima (solid dots) and maxima (open dots) over all of the different observers and texture distributions; the numbers just above them show the percentage of possible trials that each point was marked; and the horizontal bars depict the standard deviation of those judgments.
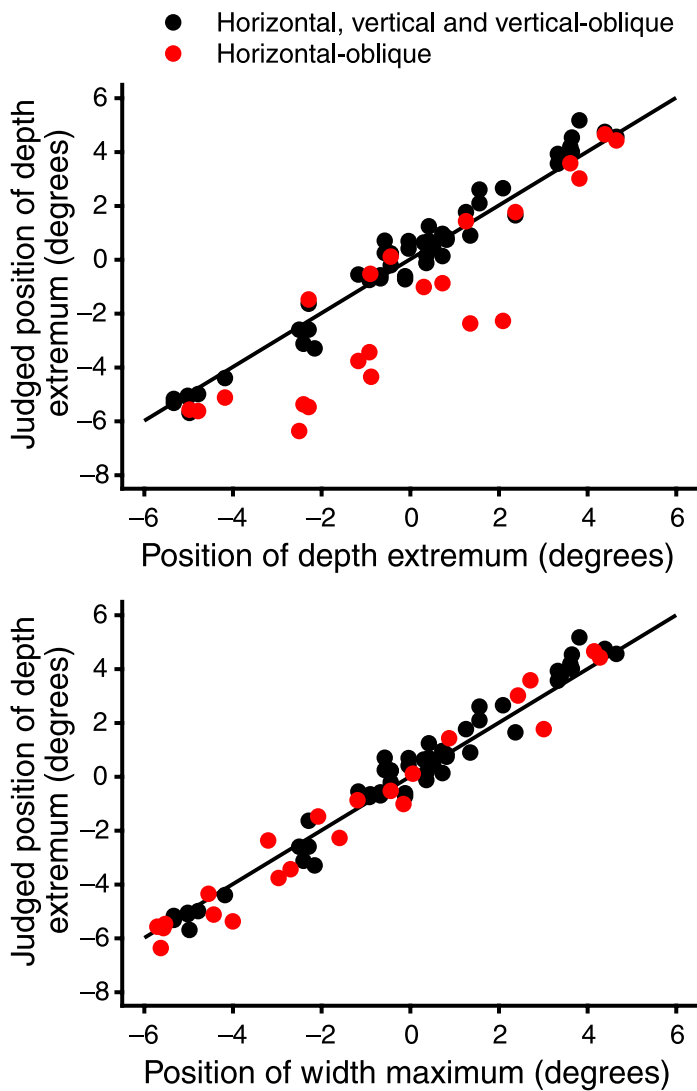
Figure 10. The average settings over all observers in Experiment 2 as a function of the ground truth (top) and the local width maxima (bottom).

any given scan line in these images provides the primary source of information for determining the ordinal pattern of perceived surface relief.

Discussion

Let us now consider how directional width gradients can be used to estimate the specific 3D shape of a surface. We shall begin with the special case of a continuously curved surface with homogeneous anisotropic texture that is viewed under orthographic projection (e.g. see Figure 2). It is important to keep in mind that there are no existing computational models for estimating shape from texture that are able to cope with this particular set of conditions, but it is possible to obtain accurate estimates of 3D shape

using the analysis of directional width gradients that was outlined in the introduction (see Figure 3).

The first stage in the computation is to interpolate the local element widths to estimate the directional width field. We suspect this will be the most difficult aspect of implementing our model in real world contexts, though some promising possibilities have been proposed by Grossberg, Kuhlmann, and Mingolla (2007) and Grossberg and Mingolla (1987). Note that this does not necessarily require the identification of individual texture elements. The underlying geometric invariant described by Equation 3 could also be estimated from the density of edges along a surface cross-section (e.g. Thaler et al., 2007), and this is likely to be a more useful measure for surfaces with more fine scale textures whose elements are not easily individuated.

Because of the effects of sampling error (or small amounts of perspective) the local width maxima along a given scan line will not all be equal to one another, which makes it difficult to determine the specific value that should be used in the denominator of Equation 3. One way to address this issue for points that are in between two different depth extrema is to use a distance weighted average of the two neighboring width maxima as the
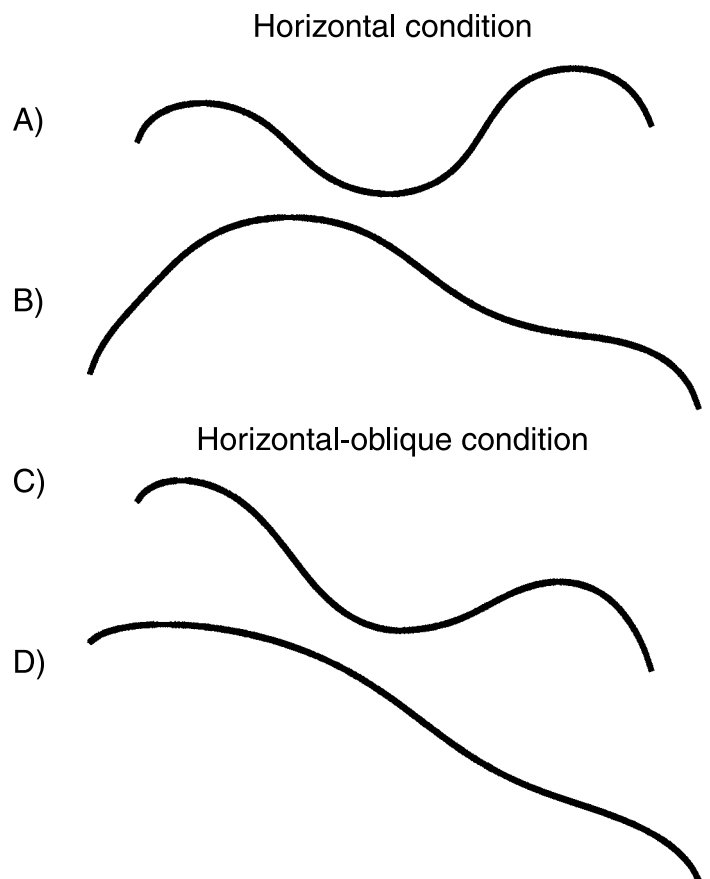

Figure 11. The patterns of surface relief that were generated based on Equation 5 for the four directional width functions depicted in Figure 9.

denominator. This guarantees that the ratio $\omega(\phi)/\omega_{\mathrm{Max}}(\phi)$ will always be exactly one at the depth extrema, and between zero and one everywhere else.

  Another important problem for estimating shapes from directional width gradients is that the sign of slant specified by Equation 3 is mathematically ambiguous. For surfaces that are bounded by smooth occlusion contours, this ambiguity can be resolved by the constraint that surface curvature perpendicular to an occlusion is always convex (Koenderink, 1984; Koenderink & Van Doorn, 1982). Thus, moving left to right from an attached smooth occlusion, the slants are initially all positive, but the sign of slant is reversed at each depth extremum that is encountered. Note, however, that the sign of slant is not reversed at inflection points, so it is important to distinguish the local width maxima at these points from those that occur at depth extrema. There are at least two sources of information for identifying inflection points. In general, the width maxima at these points will be significantly smaller than the width maxima for depth extrema. If a scan line is bounded on both ends by smooth occlusions, then it can only have an odd number of depth extrema. Thus, if there is an even number of width maxima, then at least one of them must correspond to an inflection point. A good example of this can be seen in Figure 9 for the scan lines labeled B and D. It is important to note in this figure that observers sometimes misinterpreted inflection points as depth minima, but whenever

they did so they also inserted a false alarm depth maximum so that there was always an odd number of judged depth extrema on each scan line.

  Once the width maxima $\omega_{\mathrm{Max}}(\phi)$ have been appropriately adjusted for each position ϕ along a scan line, and the width maxima that correspond to depth extrema have been distinguished from those that occur at inflection points, it is then possible to estimate the surface depth gradients using the following equation:

$$\frac{dz}{d\phi} = \tan\left(\arccos\left(\frac{\omega(\phi)}{\omega_{\mathrm{Max}}(\phi)}\right)\right). \tag{4}$$

By substituting a trigonometric identity and rearranging terms this can be reduced to:

$$\frac{dz}{d\phi} = \sqrt{\left(\frac{\omega_{\mathrm{Max}}(\phi)}{\omega(\phi)}\right)^2 - 1}. \tag{5}$$

Note that sign of slant in Equations 4 and 5 are always positive, because $\omega(\phi)/\omega_{\mathrm{Max}}(\phi)$ is constrained to be between zero and one, and they are therefore insufficient on their own for estimating 3D shape. However, this problem can easily be resolved by combining Equation 5 with the simple rule described above that the sign of slant reverses at each depth extremum. It is then possible to
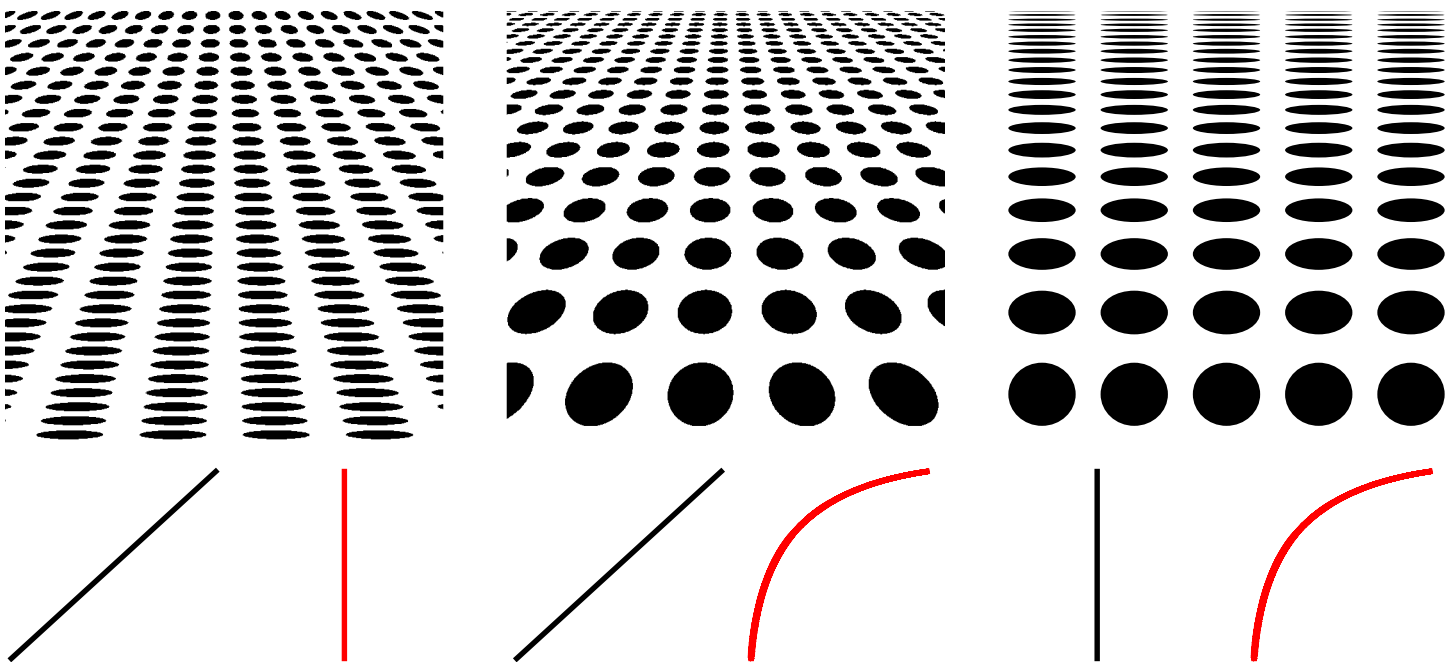

Figure 12. Three images of a planar surface. The one depicted in the center panel has a 50° slant and was rendered with a 60° field of view. The image in the left panel has the same pattern of major axis length changes as the center panel, but the widths of its texture elements in a vertical direction are all identical. The image in the right panel has the same pattern of vertical width changes as the center panel, but the major axes of all of its texture elements have identical lengths. The black and red curves below each image show the apparent depth profiles along a vertical scan line that would be predicted based on Equations 2 and 5, respectively.

estimate the surface depth profile by integrating the gradient function to determine the relative depths along each scan line.

By estimating the depth profiles along several different cross-sections in several different directions, it is a straight forward process to bind them together in a least squares sense to obtain a full blown surface reconstruction, as has been demonstrated previously by Koenderink, van Doorn, Kappers, and Todd (2001) for the analysis of observers' depth profile judgments. This is achieved by moving the profiles to and fro in depth in order to fit a surface as well as possible. This also provides a measure of goodness of fit.

Figure 11 shows the estimated patterns of surface relief that were generated with this model for the four directional width functions depicted in Figure 9. Note how the computed shapes for the scan lines (A and B) from the horizontal condition are quite similar to the ground truth (i.e. the solid black curves in Figure 9), despite the fact that the depicted object violates the model's underlying assumption about texture homogeneity. For the scan lines (C and D) from the horizontal-oblique condition, the shapes generated by the model are sheared in depth relative to the ground truth, but they correspond quite closely to how this object is perceived by human observers.

Although this model can accurately simulate the apparent shapes of continuously curved textured surfaces like elliptical cylinders or randomly deformed spheres (e.g. see Figures 2, 5, and 8), it cannot predict perceptual performance for planar or asymptotically planar surfaces. Todd et al. (2007, 2010) have shown that the perceived 3D shapes of this latter class of surfaces can be predicted quite accurately by the scaling contrast model described by Equation 2. Note that this model is fundamentally different from the analysis of directional width gradients. It computes the depth profile of a surface directly from the relative lengths of the major axes of optical texture elements (as opposed to their widths in a designated direction), and it does not involve any trigonometric transforms or integrations.

It is important to recognize that these two methods for estimating the pattern of surface relief from texture are perfect compliments of one another. Equation 2 is only effective for planar or asymptotically planar surfaces viewed with relatively large amounts of perspective, because those are the only conditions that produce significant variations in texture scaling. The analysis of directional width gradients, on the other hand, is only effective for continuously curved surfaces viewed with minimal perspective.

To better appreciate how these models could be used in conjunction with one another, it is useful to consider the three images presented in Figure 12. The surface depicted in the center panel has a 50° slant and was rendered under perspective projection with a 60° field of view. The image in the left panel has the same pattern of major axis length changes as the center panel, but the widths of its texture

elements in a vertical direction are all identical. Conversely, the image in the right panel has the same pattern of vertical width changes as the center panel, but the major axes of all of its texture elements have identical lengths. The black and red curves below each image show the apparent depth profiles along a vertical scan line that would be predicted based on Equations 2 and 5, respectively. It is clear that neither model alone can predict the apparent surface structure from all three of
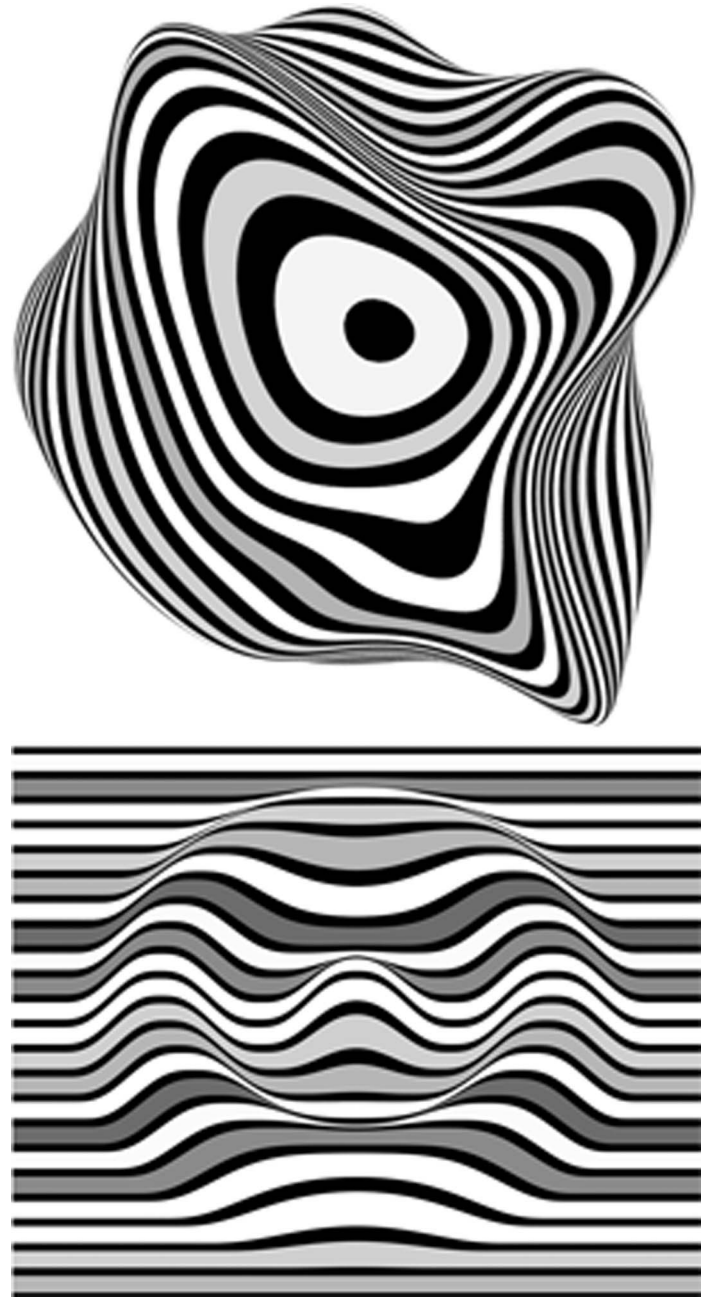

Figure 13. Two images of surfaces with contour textures whose perceptual appearance cannot be predicted from the models described by Equations 2 and 5.

these displays. For the left and middle images with systematic variations in texture scaling, observers' perceptions are quite consistent with Equation 2, just as they were for the images of hyperbolic cylinders used by Todd et al. (2007). However, for the image on the right with no variation in texture scaling, observers' perceptions are better described by an analysis of directional width gradients, just as they were in the present experiments for images of ellipsoids or randomly deformed spheres (see also Todd & Akerstrom, 1987).

These observations suggest that the perception of 3D shape from texture may involve an analysis that is context dependent (see also Cutting & Millard, 1984). When there are significant variations of texture scaling, observers' perceptions will be primarily determined by gradients of texture length, as described by Equation 2. When this information is unavailable, however, observers will then rely on directional width gradients, based on Equation 5. This two-pronged strategy is ecologically quite sensible, because it facilitates the perception of 3D shape from texture over a wider range of conditions than would be possible using either of these models in isolation.

Although Equations 2 and 5 can account for a wide range of phenomena, a complete theoretical explanation of the perception of 3D shape from texture is likely to require at least one more context dependent module in order to deal with textures composed of extended contours or orientation flows. Figure 13 shows two examples of how patterns of image contours can produce compelling impressions of 3D surfaces. One possible approach for estimating shape from contours is to constrain the problem by assuming that the depicted surface is developable (i.e. that it has no curvature in one direction), and that its contours are all oriented in the direction of maximum curvature (Stevens, 1981b) or along surface geodesics (Knill, 1992, 2001). The empirical evidence suggests, however, that these models cannot account for the apparent shapes of surfaces from contour textures. For example, observers are able to make accurate shape judgments for the images in Figure 13 (see Todd et al., 2004; Todd & Reichel, 1990), despite the fact that the contours on the depicted surfaces do not come close to satisfying any of the models' underlying assumptions. How observers are able to achieve this high level of performance is an interesting problem that will remain for future research.

Acknowledgments

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

References

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

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/content/6/9/7, doi:10.1167/6.9.7. [PubMed] [Article]

Bajcsy, 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. [Article]

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. [Article]

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. [Article]

Cutting, J. E., & Millard, R. T. (1984). Three gradients and the perception of flat and curved surfaces. *Journal of Experimental Psychology: General, 113,* 198–216. [PubMed]

Fleming, R. W., Torralba, A., & Adelson, E. H. (2004). Specular reflections and the perception of shape. *Journal of Vision, 4*(9):10, 798–820, http://journalofvision.org/content/4/9/10, doi:10.1167/4.9.10. [PubMed] [Article]

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

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

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

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

Grossberg, S., Kuhlmann, L., & Mingolla, E. (2007). A neural model of 3D shape-from texture: Multiple-scale

filtering, boundary grouping, and surface filling-in. *Vision Research, 47,* 634–72.

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. [Article]

Knill, D. C. (1992). Perception of surface contours and surface shape: From computation to psychophysics. *Journal of the Optical Society of America A, Optics and Image Science, 9,* 1449–1464. [PubMed] [Article]

Knill, D. C. (2001). Contour into texture: Information content of surface contours and texture flow. *Journal of the Optical Society of America A, Optics, Image Science, and Vision, 18,* 12–35. [PubMed]

Koenderink, J. J. (1984). What does the occluding contour tell us about solid shape? *Perception, 13,* 321–330. [PubMed]

Koenderink, J. J., & van Doorn, A. J. (1982). The shape of smooth objects and the way contours end. *Perception, 11,* 129–137. [PubMed]

Koenderink, J. J., van Doorn, A. J., Kappers, A. M. L., & Todd, J. T. (2001). Ambiguity and the "Mental Eye" in pictorial relief. *Perception, 30,* 431–448. [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. [Article]

Malik, J., & Rosenholtz, R. (1994). A computational model for shape from texture. In G. R. Bock & J. A. Goode (Eds.), *Higher order processing in the visual system* (pp. 272–286). Chichester: John Wiley & Sons.

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

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

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

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]

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

Stevens, K. A. (1981b). The visual interpretation of surface contours. *Artificial Intelligence,* 17, 47–53.

Stevens, K. A. (1984). On gradients and "texture" gradients. *Journal of Experimental Psychology: General, 113,* 217–224. [PubMed]

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

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

Todd, J. T., & Akerstrom, R. A. (1987). The perception of three dimensional form from patterns of optical texture. *Journal of Experimental Psychology: Human Perception and Performance, 13,* 242–255. [PubMed] [Article]

Todd, J. T., Christensen, J. C., & Guckes, K. M. (2010). Are discrimination thresholds a valid measure of variance for judgments of slant from texture [Abstract]? *Journal of Vision, 10*(2):20, 1–18, http://journalofvision.org/content/10/2/20, doi:10.1167/10.2.20.

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] [Article]

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

Todd, J. T., & Reichel, F. D. (1990). The visual perception of smoothly curved surfaces from double projected contour patterns. *Journal of Experimental Psychology: Human Perception and Performance, 16,* 665–674. [PubMed] [Article]

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

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 3D shape from texture. *Journal of Vision, 7*(12):9, 1–16, http://journalofvision.org/content/7/12/9, doi:10.1167/7.12.9. [PubMed] [Article]

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