The perception of 3D shape from planar cut contours

Eric J. L. Egan  
Department of Psychology, The Ohio State University, Columbus, OH, USA

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

Flip Phillips  
Skidmore College, USA

A new computational analysis is described for estimating 3D shapes from orthographic images of surfaces that are textured with planar cut contours. For any given contour pattern, this model provides a family of possible interpretations that are all related by affine scaling and shearing transformations in depth, depending on the specific values of its free parameters that are used to compute the shape estimate. Two psychophysical experiments were performed in an effort to compare the model predictions with observers’ judgments of 3D shape for developable and non-developable surfaces. The results reveal that observers’ perceptions can be systematically distorted by affine scaling and shearing transformations in depth and that the magnitude and direction of these distortions vary systematically with the 3D orientations of the contour planes.

Keywords: 3D surface and shape perception, shape and contour, depth, ecological optics, computational modeling, texture


Introduction

A popular method for depicting the 3D structure of a curved surface involves displaying a pattern of contours formed by the intersection of the surface with a series of parallel planes (e.g., see Figure 1). This technique was first developed in the field of cartography (Robinson & Thrower, 1957; Tanaka, 1932), where it is sometimes referred to as the method of inclined planes. It is also used in mechanical drawing and computer graphics for the visualization of 3D surfaces (Williamson, 1971; Wright, 1973), and the patterns it produces are quite similar to those that have been created by op artists, such as Josef Albers, Frank Stella, Bridget Riley, and Victor Vasarely. In the natural environment, planar cut contours are frequently observed in the layered strata of rock formations, terraced landscapes, and man-made objects composed of layered materials such as butcher block. Although there have been numerous psychophysical investigations that have examined observers’ perceptions of surfaces depicted with planar cut contours (Bocheva, 2009; Li & Zaidi, 2000, 2004; Mamassian & Landy, 2001; Shapley & Maertens, 2008; Stevens & Brooks, 1987; Todd & Oomes, 2002; Todd, Oomes, Koenderink, & Kappers, 2004; Todd & Reichel, 1990; Todd, Thaler, & Dijkstra, 2005; Todd, Thaler, Dijkstra, Koenderink, & Kappers, 2007; Tse, 2002), there is currently no consensus about why this technique is so perceptually effective or how such patterns are able to specify the 3D structure of a surface.

The first computational analysis of shape from contour textures was performed by Stevens (1981, 1986). His analysis was based on three critical assumptions: First, that the underlying surface is developable (i.e., that one of its principal curvatures is zero); second, that the surface contours are lines of maximum curvature; and third, that the images to be analyzed are rendered with negligible perspective. Whenever these conditions are satisfied, the local surface normal (N) at any point along a contour can be computed from the following equation:

$$n = \left\{ -a \sin \beta, \cos \beta \left( \frac{a^2 + 1}{a} \right), \sin \beta \right\},$$  (1)

where $\beta$ is the 2D angle between the 2D ruling vector and a vector tangent to the 2D contour, and $a$ is the $z$-component of the surface ruling vector (i.e., the direction of zero curvature). Note that $\beta$ can be measured from 2D image data, whereas $a$ is an unknown free parameter. It is important to recognize that lines of curvature on a developable surface are also planar cuts, so that the set of images for which Stevens’ analysis is applicable constitutes a subset of those that can be generated using the method of inclined planes.

A similar analysis for computing shape from contours was later developed by Knill (1992, 2001), which shares two of the key assumptions of Stevens’ model—namely, that the underlying surface is developable and that it is viewed with negligible perspective. With respect to the surface contours, however, Knill’s analysis is based on a
Figure 1. A pattern of image contours that is perceived as a smoothly curved 3D surface.

less restrictive assumption that the contours are geodesics rather than lines of curvature. Whenever these assumptions are satisfied, the local surface normal ($N_S(s)$) at any point along a contour can be computed using the following differential equation:

$$
\frac{\partial N_S}{\partial s} = k(s) \frac{\{[r(s) \wedge V]N_S(s)\} \wedge N_S(s)}{|[r(s) \wedge V] \wedge N_S(s)|} \\
\times \frac{|[n(s) \wedge N_S(s)] \wedge V|}{|[N_S(s) \wedge V, n(s)]|} \\
\times \frac{1}{\{1 - \left(\frac{[r(s) \wedge V] \wedge N_S(s)}{|[r(s) \wedge V] \wedge N_S(s)|}, \frac{m(s) \wedge N_S(s)}{|m(s) \wedge N_S(s)|}\right)^2\}^2}
$$

(2)

where $\kappa(s)$ is the curvature of the contour parameterized in terms of arc length, $n(s)$ is the 2D normal vector to the contour parameterized in terms of arc length, $r(s)$ is the 2D surface ruling parameterized in terms of arc length, and $V$ is the viewing direction. In order to solve this equation, it is necessary to specify the surface normal at one point along the contour to provide the initial conditions. Because the slant and tilt of this normal are unknown, the estimated shape is ambiguous up to a 2-parameter family of possible solutions. Although geodesics on a surface are generally not planar cuts (except along lines of curvature), their optical projections are often quite similar, especially for surfaces that are slanted in depth. Thus, this analysis could potentially be used to obtain a qualitatively correct interpretation of 3D shape for planar cut contours on developable surfaces (Knill, 1992).

An important limitation of both the Stevens and Knill models is that they are only applicable to surfaces that are developable, and they cannot therefore explain the compelling perceptions of 3D shape that occur for images of non-developable surfaces that have been created using the method of inclined planes (e.g., see Figure 1). The most obvious approach for analyzing such images is to exploit the constraint that the surface contours are planar cuts. The earliest attempts to incorporate this assumption were developed in computer vision for contour patterns on generalized cylinders (Horaud & Brady, 1988; Ulupinar & Nevatia, 1995). In the discussion that follows, we will describe a more general analysis that is applicable to arbitrary curved surfaces.

Consider the orthographic projection of a pattern of surface contours created from inclined planes with a slant $\sigma$ and a tilt $\tau$. Let us begin by constructing an index to represent the order of the contour planes and proportionally subdividing the distances between them to obtain a continuous scale in texture space. This makes it possible to assign a unique value ($v$) of the texture scale to any given point $(x, y)$ in an image based on the particular contours in its immediate local neighborhood. For convenience, we will align the $y$-axis of the image plane coordinate system to be parallel to the tilt of the contour planes. From the texture scale value at each point on the surface, we can compute the relative depth of that point from the following equation:

$$
Z = \frac{Sv + y\sin\sigma}{\cos\sigma},
$$

(3)

where $S$ is a scaling factor that defines how the texture index is related to distances in physical space. It is important to recognize that $\sigma$, $\tau$, and $S$ are unknown free parameters. Thus, this analysis can determine the relative pattern of depth on a surface up to a three-parameter family of possible interpretations, without requiring any assumptions about the underlying surface geometry. There is one degenerate case, however, that deserves to be highlighted. If the slants of the contour planes are close to $\pm 90^\circ$, then Equation 3 cannot be evaluated. The image contours in that case are reduced to a pattern of parallel straight lines that provide no useful information about 3D shape.

It is interesting to note that none of the analyses described above are able to provide a unique interpretation of 3D shape from surface contours. Instead, they all allow a family of possible interpretations, much like the analyses of structure from motion or shape from shading (see Todd, 2004). Whenever their underlying assumptions are satisfied, the ambiguous interpretations for all three of these models can be related to one another by an affine shearing transformation in depth. For Stevens’ model, this shear is restricted to a single direction, whereas it can be oriented in any direction for the other two. The planar cut model also involves an ambiguous depth scaling.

How might observers resolve this ambiguity in order to achieve a stable perception? In some cases, there is information available within an image to specify one or
more of the unknown free parameters. With planar cut contours, for example, the tilts of the contour planes can often be estimated from their global orientations or foreshortening within the image plane. Another popular strategy is to select an interpretation that minimizes some aspect of a surface’s 3D structure. For example, Stevens (1981) proposed resolving the ambiguity of his model by selecting a solution that minimizes the overall surface slant.

**Experiment 1**

Given the theoretical ambiguity of surface contours, it would not be surprising if observers’ judgments of 3D shape from contours vary systematically as a function of how closely the assumed values of the free parameters conform to the ground truth. Figure 2 provides some anecdotal evidence that this is indeed the case. The images in this figure all depict the same 3D surface, but they differ with respect to the relative orientations of the contour planes. Note how this produces a systematic variation in their apparent 3D structures. The research described in Experiment 1 was designed to measure the specific shapes that observers perceive from these images and to assess whether those perceived shapes are consistent with the families of possible interpretations for the three alternative models described above.

**Methods**

**Subjects**

Seven observers participated in the experiment, including one of the authors (JT) and six others who were naive about the issues being investigated. All of the observers had normal or corrected-to-normal visual acuity, and they all wore an eye patch to eliminate conflicting flatness cues from binocular vision.

**Apparatus**

The experiment was conducted using a Dell Dimension 8300 PC with an ATI Radeon 9800 PRO graphics card and a 19-in CRT with a spatial resolution of 1280 × 1024 pixels. The stimulus images were presented within a 25.4 × 19.0 cm region (800 × 600 pixels) of the display screen, which subtended 14.6° × 10.9° of visual angle when viewed at a distance of 100 cm. A second LCD monitor was located directly to the left of the CRT that contained a series of adjustable dots that observers manipulated to record their responses. A chin rest was used to restrict head movements.

**Stimuli**

Images of a corrugated surface were rendered using 3D Studio Max by Autodesk with procedural contour textures. The surface was rendered under orthographic projection in the same fixed position for all stimuli. The
surface had a global slant of $42.7^\circ$ and a global tilt of $25^\circ$.
The textures were created by a volume of alternating light and dark rectangular slabs that produced a series of parallel planar cuts through the surface. The inclined planes had slants and tilts of $(-63^\circ, 54^\circ)$, $(-53^\circ, 37^\circ)$, $(-49^\circ, 19^\circ)$, $(-47^\circ, 0^\circ)$, and $(-49^\circ, -19^\circ)$, respectively, for each of the five stimulus conditions. In an object-based frame of reference, the inclined planes were all perpendicular to the global orientation of the depicted surface, and they were tilted by $12.5^\circ$, $-12.5^\circ$, $-25^\circ$, and $-37.5^\circ$, respectively, relative to the direction of maximum curvature. Five images with random white/gray color schemes were rendered for each possible contour orientation to create twenty-five total stimuli. An example of each condition is shown in Figure 2. Stimulus condition B is a special case where the contours are both geodesics and lines of maximum curvature.

**Procedure**

A profile reproduction task was employed in order to measure perceived relative depth between multiple surface locations (see Koenderink, van Doorn, Kappers, & Todd, 2001). On each trial, a stimulus was presented on the CRT and one of its horizontal scan lines was marked by a row of nine equally spaced dots. On the second monitor, an identical row of dots was presented on a black background (see Figure 3). A mouse was used to drag each of the dots on the second monitor up and down. Observers were asked to adjust the dots on the second monitor in order to match the apparent surface profile in depth of the dots superimposed on the stimulus. When they were satisfied with their adjustments, they initiated a new trial by pressing the space bar.

The dots were superimposed onto the stimuli on one of five horizontal scan lines, which were vertically aligned and evenly distributed so that each consecutive pair of scan lines was separated by 100 vertical image pixels. Responses were obtained for each possible scan line in each condition in a random order. The dots had diameters of fifteen pixels, and they were horizontally separated by 76 pixels. At this spacing, seven contiguous markers covered one wavelength of the corrugated surface. The dots of each scan line corresponded to the same six phase intervals with respect to the corrugated surface. Thus, even though there were 45 different probe dot locations in each condition, they all could be mapped to one of the six possible phases. All of the dots were red except the center dot, which was green. This was done to create an easily identifiable reference point when switching between displays.

The instructions given to each observer emphasized the importance of local depth differences (i.e., consecutive dots) as well as global differences (e.g., the first dot compared to the last). Each observer completed five to ten practice trials with an experimenter present until both parties were confident that the observer understood the task. The experiment was organized into five blocks, each of which included one presentation for each of the five possible scan lines for each of the five stimulus categories. The mean time per trial was between 1 and 2 min. All observers completed the experiment within a 30- to 45-min time period.

**Results**

For each observer in each condition, the average judged depth difference was calculated between each pair of probe dots on all of the different scan lines. Because all of these scan lines had the same six phases of the surface depth profile, it was possible to collapse the judgments for

![Figure 3](image-url)
different scan lines into a single apparent profile in depth. In order to assess the reliability of these judgments, an analysis of linear regression was performed to compare the responses obtained in blocks 2 and 3 of the experiment for each observer with those produced in blocks 4 and 5. The results revealed that the test–retest reliability for this task was quite high, such that the average coefficient of determination ($R^2$) was 0.84. A similar analysis was used to measure the consistency among all possible pairs of observers, and the average value of $R^2$ for these correlations was 0.79.

The results obtained for each of the different contour orientations are presented in Figure 4, which shows the average response profile over all observers in each condition (red circles) together with the ground truth (black curves). Error bars are not shown because most of them are smaller than the radii of the red circles. Note that the response profiles are systematically underestimated, and in many of the conditions, they are also sheared relative to the ground truth (cf. Koenderink et al., 2001). For example, the mean response for Condition A was sheared, such that the surface point corresponding to the leftmost peak of the response profile appeared closer in depth than the one corresponding to the rightmost peak.

The opposite is true for Condition E. An analysis of variance revealed that the effect of contour plane orientation was statistically significant, $F(4, 24) = 12.13$, $p < 0.001$. Affine correlations between the judged depth profiles and the ground truth were performed to measure the magnitude of these perceptual distortions in each condition using the following linear model:

$$Z' = X_{shear} \times X + Z_{scale} \times Z,$$

where $Z'$ is the judged depth of a given probe point, $X$ is its horizontal position along the scan line, $Z$ is the true depth of the point in physical space, and $(X_{shear}, Z_{scale})$ are the affine coefficients to be estimated. The results of this analysis are presented in Table 1. If observers’ perceptions had been based on the analyses proposed by Stevens (1981, 1986) or Knill (1992, 2001), then their judgments should have been most accurate in Condition B, because that is the only condition for which the contours were aligned along surface geodesics or lines of curvature. However, the results do not confirm this prediction. The judged depth profiles in Condition D had the least amount of shear relative to the ground truth. That is the condition in which the horizontal scan line was perpendicular to the tilt of the contour planes.

We also performed affine correlations on the judgments obtained from individual observers, and the results, on average, were nearly identical to those obtained from the composite data described in Table 1. The standard deviations of the $X_{shear}$ and $Z_{scale}$ parameters among different observers averaged over conditions were 0.09 and 0.14, respectively.

It is possible that uncontrolled 2D cues such as the absence of accommodative blur (Watt, Akeley, Ernst, & Banks, 2005) may have contributed to the underestimation of surface depth in these displays, although the effect of these cues is typically much smaller than the 47.4% underestimation that was evident in the observers’ judgments. It should also be noted in this regard that the presence of 2D cues cannot be responsible for the shearing distortions in the judged patterns of relief or for the significant differences among the five orientation conditions.

Additional analyses were performed to determine how closely the three models described in the introduction could reproduce the ground truth in the different conditions and whether they could account for the systematic distortions in

<table>
<thead>
<tr>
<th>Condition</th>
<th>$R^2$</th>
<th>$X_{shear}$</th>
<th>$Z_{scale}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.98</td>
<td>−0.18</td>
<td>0.56</td>
</tr>
<tr>
<td>B</td>
<td>0.99</td>
<td>−0.13</td>
<td>0.54</td>
</tr>
<tr>
<td>C</td>
<td>0.99</td>
<td>−0.07</td>
<td>0.54</td>
</tr>
<tr>
<td>D</td>
<td>0.98</td>
<td>−0.03</td>
<td>0.53</td>
</tr>
<tr>
<td>E</td>
<td>0.93</td>
<td>0.14</td>
<td>0.47</td>
</tr>
</tbody>
</table>

Table 1. Affine correlations between the observers’ judgments and the ground truth for each condition of Experiment 1.
the behavioral data. This was achieved by implementing each model for the five different contour patterns and computing the values of their free parameters that provide the best least-squares fits to the ground truth and to the average response profiles. The model implementations all began using traditional morphological techniques to binarize and extract local edges as pixel locations. The orientation and curvature functions were computed directly from these representations. All remain-
ing contours in an image were generated from the longest one by translating it along the ruling direction. The NDSolve[] differential equation solver from Mathematica was used to evaluate Equation 2. Because the analyses proposed by Stevens and Knill compute local surface normals along a contour, it was necessary to compute the local depth gradients from the normals, and then integrate the gradient function along a scan line in order to determine the predicted depth profile for any given set of parameter values.

A grid search was performed over the entire parameter space for all three models in order to identify the specific values that provided the best fits to the data in all of the different conditions. These fits were performed on the judgments of each individual observer and also on the composite depth profiles obtained by averaging over

<table>
<thead>
<tr>
<th>Condition</th>
<th>Fit to ground truth</th>
<th>Fit to data</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$R^2$</td>
<td>Slope</td>
</tr>
<tr>
<td><strong>Stevens’ model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0.93</td>
<td>3.24</td>
</tr>
<tr>
<td>B</td>
<td>0.99</td>
<td>1.01</td>
</tr>
<tr>
<td>C</td>
<td>0.99</td>
<td>1.46</td>
</tr>
<tr>
<td>D</td>
<td>0.99</td>
<td>11.50</td>
</tr>
<tr>
<td>E</td>
<td>0.58</td>
<td>66.67</td>
</tr>
<tr>
<td><strong>Knill’s model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0.99</td>
<td>0.88</td>
</tr>
<tr>
<td>B</td>
<td>0.99</td>
<td>1.01</td>
</tr>
<tr>
<td>C</td>
<td>0.99</td>
<td>0.93</td>
</tr>
<tr>
<td>D</td>
<td>0.99</td>
<td>0.89</td>
</tr>
<tr>
<td>E</td>
<td>0.99</td>
<td>0.88</td>
</tr>
<tr>
<td><strong>Planar cut model</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0.99</td>
<td>1.02</td>
</tr>
<tr>
<td>B</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>C</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>D</td>
<td>0.99</td>
<td>1.00</td>
</tr>
<tr>
<td>E</td>
<td>0.99</td>
<td>1.00</td>
</tr>
</tbody>
</table>

Table 2. Fits to the ground truth, fits to the observer’s judgments, and the Akaike information criterion (AIC) for the Stevens, Knill, and planar cut models in each condition of Experiment 1.

observers. The average $R^2$ for the individual fits over all observers, models, and conditions was 0.89, which is somewhat noisier than those obtained from the composite data that had an average $R^2$ of 0.94. In all other respects, however, the individual and composite fits revealed the same basic pattern of results.

A complete summary of the composite fits is provided in Table 2. The two leftmost columns of numbers show the linear correlations and slopes of regression between the ground truth and the best-fitting profiles generated by the models for each of the different contour orientations. As expected, the Stevens and Knill models provide excellent fits to the ground truth in Condition B, which satisfies their underlying assumptions, and the accuracy of the fits is diminished with increasing violations of these assumptions. Also to be expected is that the planar cut model provides excellent fits to the ground truth in all conditions, because all of the images were generated with planar cut textures.

The three rightmost columns of numbers in Table 2 show the linear correlations and slopes of regression between the observers’ judgments and the best-fitting profiles generated by the models and the Akaike information criterion (AIC) for each of the three models with each possible orientation of the contour planes. Note that the Stevens and Knill models do poorly in all conditions. Although they can account for the shear of the profiles, they have no mechanism to couple this with the correct amount of depth scaling. Both of these models predict that the apparent depth scaling should covary with the magnitude of apparent shear but that prediction is not confirmed by the observers’ judgments.

The planar cut model, in contrast, provides excellent fits to the data in all conditions. Table 3 shows the best-fitting parameter values for this model from the composite depth profiles in each of the five conditions. The same basic pattern of results is also revealed by the fits obtained from individual observers. The standard deviations of the $S$, $\sigma$, and $\tau$ parameters among different observers averaged over conditions were 0.13, 10.40°, and 3.98°, respectively.

Given the systematic variation of the $\sigma$ and $\tau$ parameters across conditions, it is unlikely that these were determined by statistical priors. Rather, these results suggest that there is visual information within the 2D images to estimate the orientations of the contour planes. One likely source of information about the tilts of these

<table>
<thead>
<tr>
<th>Condition</th>
<th>$S$</th>
<th>$\sigma$</th>
<th>$\tau$</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.60</td>
<td>$-60$</td>
<td>25</td>
</tr>
<tr>
<td>B</td>
<td>0.53</td>
<td>$-57$</td>
<td>14</td>
</tr>
<tr>
<td>C</td>
<td>0.52</td>
<td>$-44$</td>
<td>10</td>
</tr>
<tr>
<td>D</td>
<td>0.58</td>
<td>$-37$</td>
<td>1</td>
</tr>
<tr>
<td>E</td>
<td>0.47</td>
<td>$-44$</td>
<td>5</td>
</tr>
</tbody>
</table>

Table 3. The best-fitting parameter values for the planar cut model in each condition of Experiment 1.
planes is provided by the orientations of the 2D contours. Indeed, the best-fitting tilts are almost perfectly correlated with the direction of the average 2D contour normal within each image \((R = 0.99)\). There is also a high negative correlation between the best-fitting slants and the amplitudes of the 2D contours \((R = -0.88)\). The reason for this latter effect, we suspect, is that the apparent slant of a surface increases with the 2D amplitude of its projected contours and that the estimated slants of the contour planes are negatively related to the apparent surface slants.

In one respect, it is perhaps not surprising that the planar cut model produces the best fits, because it has the largest number of free parameters. Indeed, if an additional depth scaling parameter were added to the Stevens or Knill models, their fits would be almost as good as those produced by the planar cut model. The problem with that approach, however, is that it has no theoretical motivation except as a post hoc kluge. It would also result in an internal inconsistency. It is important to keep in mind that the free parameters of the Stevens and Knill models specify the 3D orientation of a point on the depicted surface. If a depth scaling parameter were added, then the orientation of that point on the estimated surface could be dramatically different from the parameter values that were used to create that estimate.

### Experiment 2

One of the primary advantages of the planar cut model for determining shape from contours is that it has no constraints on the underlying surface geometry. This is in stark contrast to the earlier approaches of Stevens (1981, 1986) and Knill (1992, 2001), which are only applicable to developable surfaces—i.e., those for which one of the principal curvatures at every point is zero. In an effort to test the psychological validity of this constraint, Experiment 2 was designed to investigate observers’ judgments of shape from contours for non-developable surfaces.

### Methods

The apparatus and procedure were identical to those described for Experiment 1. The stimuli included images of two randomly deformed spheres (see Figure 5) and each of these objects was textured with a series of parallel planar cuts in two different orientations. For object O1, the contour planes had possible orientations (slant, tilt) of \((30^\circ, 0^\circ)\) and \((30^\circ, 90^\circ)\) in a polar north coordinate system. For Object O2, the possible orientations were \((30^\circ, 180^\circ)\) and \((0^\circ, 0^\circ)\). On each trial, a stimulus was presented on the CRT together with a collinear set of nine equally spaced dots. On the second monitor, an identical row of dots was presented on a black background. Observers were asked to adjust the dots on the second monitor in order to match the apparent surface profile in depth of the dots superimposed on the stimulus. Unlike Experiment 1, these dots could appear in either a horizontal or a vertical orientation, and each horizontal scan line had one dot in common with each vertical scan line, whose locations are identified by the flanking red lines in Figure 5. Each of the four scan lines on each of the four stimuli was judged four times during a single experimental session. The displays were judged by six observers, including two of the authors (EE and JT) and four others who were naive about how the displays were created.

### Results

For each observer in each condition, the average judged depth difference was calculated between each pair of

![Figure 5](image-url)
Table 4. Affine correlations between the observers’ judgments and the ground truth for each condition of Experiment 2.

<table>
<thead>
<tr>
<th>Object ($\sigma$, $\tau$)</th>
<th>$R^2$</th>
<th>Xshear</th>
<th>Yshear</th>
<th>Zscale</th>
</tr>
</thead>
<tbody>
<tr>
<td>O1 ($30^\circ$, $0^\circ$)</td>
<td>0.92</td>
<td>0.02</td>
<td>0.39</td>
<td>0.76</td>
</tr>
<tr>
<td>O1 ($30^\circ$, $90^\circ$)</td>
<td>0.92</td>
<td>0.29</td>
<td>0.03</td>
<td>0.83</td>
</tr>
<tr>
<td>O2 ($30^\circ$, $180^\circ$)</td>
<td>0.92</td>
<td>0.00</td>
<td>-0.39</td>
<td>0.72</td>
</tr>
<tr>
<td>O2 ($0^\circ$, $0^\circ$)</td>
<td>0.93</td>
<td>0.03</td>
<td>-0.01</td>
<td>0.78</td>
</tr>
</tbody>
</table>

(Where $Z$ is presented in Table 4. We also performed affine correlations on the judgments obtained in the first half of the experiment for each observer with those produced in the second half. The results revealed that the test–retest reliability for this task was quite high, such that the average $R^2$ was 0.89. A similar analysis was used to measure the consistency among all possible pairs of observers, and the average $R^2$ for these correlations was 0.73.)

As in Experiment 1, the judgments of all six observers were combined to compute an average depth profile for each of the possible scan lines in each condition. In order to measure any systematic perceptual distortions in these judgments, it was first necessary to adjust the different scan lines in depth in order to obtain a maximally smooth surface (see Koenderink et al., 2001). This was achieved by minimizing the least-squares difference between the four pairs of points that overlapped one another on different scan lines using the LeastSquares[] function in Mathematica. An affine correlation was then performed to compare the judged patterns of relief with the ground truth using the following linear model:

$$Z' = X\text{shear} \times X + Y\text{shear} \times Y + Z\text{scale} \times Z,$$

where $Z'$ is the judged depth of a given probe point, $(X, Y,$ and $Z)$ are the Cartesian coordinates of that point in physical space, and $(X\text{shear}, Y\text{shear}, Z\text{scale})$ are the affine coefficients to be estimated. The results of this analysis are presented in Table 4. We also performed affine correlations on the judgments obtained from individual observers. The average $R^2$ values for the individual correlations were slightly lower than those for the composite data (0.86 versus 0.92), but the relative pattern of coefficients in the different conditions was nearly identical. These findings confirm the theoretical prediction of the planar cut model that the possible interpretations for any given contour pattern are related to one another by affine scaling and shearing transformations in depth.

It is important to keep in mind, however, that the planar cut model is more constrained than the affine transformation described in Equation 5, because the effects of altering its three free parameters ($S$, $\sigma$, and $\tau$) are not independent—i.e., they can all have a coordinated influence on both shear and depth scaling. Thus, an additional analysis was performed to determine how closely the planar cut model could account for the specific patterns of distortion in these data. This was achieved by implementing the model for the four different contour patterns and computing the values of its free parameters that provide the best least-squares fits to the average response profiles. The results of this analysis (see Table 5) reveal that the planar cut model can account, on average, for 85% of the variance in the perceived relative depths between the different probe points in each condition, and similarly good fits were also obtained from analyses of the individual observers.

There is an interesting pattern in these data that may not be obvious in Tables 4 and 5 but deserves to be highlighted nonetheless. The black dots in Figure 6 depict the 3D orientations of the contour planes in each condition in a polar north coordinate system, where slant is represented by the radial component (i.e., the distance of a dot from the center), and tilt is represented by the angular component (i.e., the clock position). The green dots in this figure show the estimated slants and tilts from the planar cut model that produced the best fits to the data in each condition. Note that in the three conditions with slanted contour planes, the estimated tilts are quite accurate but that the estimated slants are much smaller than the ground truth. Finally, the red dots in Figure 6 show the slants and tilts of a fronto-parallel plane after it is transformed by the best-fitting shear parameters in each condition that are listed in Table 3. Note that these shearings deformations are all also closely aligned with the tilts of the contour planes. These findings suggest that there is optical information from the 2D orientations of the image contours to specify the tilts of the contour planes but that there is no sufficient information to pin down the slant or the texture scaling. When these parameters are underestimated for the computation of 3D shape using a planar cut analysis, the slant of the contour planes is misinterpreted as a global slant of the depicted surface, and a similar misinterpretation is evident in the judgments of human observers as shown in Figure 6.

### Discussion

The research described in the present article was designed to investigate how human observers determine
the 3D shape of an object from patterns of image contours formed by parallel planar cuts through its surface. In Experiment 1, the stimuli depicted corrugated developable surfaces with contour planes in several different orientations. In Experiment 2, the stimuli depicted non-developable surfaces that were created by random deformations of a sphere. The results revealed that observers’ shape judgments for both types of surfaces are systematically distorted relative to the ground truth by affine shearing and scaling transformations in depth and that the precise manner of these distortions varies with the 3D orientations of the contour planes that were used to generate each texture.

We have also presented a new computational analysis for estimating 3D shapes from orthographic images of surfaces with planar cut contours. For any given contour pattern, this model provides a family of possible interpretations that are all related by affine scaling and shearing transformations in depth, depending on the specific values of the three free parameters that are used to compute the shape estimate. Although this ambiguity may appear at first blush to be a weakness of the model, it is remarkably consistent with the distortions of judged shape that are exhibited by human observers. Indeed, the overall pattern of systematic errors is quite similar to what would be expected if observers had determined the 3D shapes of the objects using a planar cut analysis but with incorrect estimates of the slant and scaling of the contour planes.

One important advantage of our planar cut analysis is that it imposes no restrictions on the underlying surface geometry. This stands in stark contrast to the methods proposed by Knill (1992, 2001) and Stevens (1981, 1986), which are only applicable to developable surfaces—i.e., those for which one of the principal curvatures at every point is zero. The evidence is quite clear, however, that images of non-developable surfaces with planar cut contours produce compelling perceptions of 3D shape (Bocheva, 2009; Todd & Oomes, 2002; Todd et al., 2004; Todd & Reichel, 1990; Tse, 2002). The analysis described by Equation 3 is the only one proposed to date that could potentially explain how this is possible.

One of the central components of our planar cut analysis is that the contours must be indexed in order to assign a unique value \( v \) of the texture scale to any given point \( (x, y) \) in an image based on the particular contours in its immediate local neighborhood. The discussion thus far has implied that the contour planes must be equally spaced to create a uniform texture scale, but this is clearly too strong an assumption. Indeed, the stimuli employed in Experiment 1 were all generated with contour planes that were unequally spaced. Note in Figure 2, for example, that obtain a depth profile from that, it is necessary to compute the local depth gradients from the normals and then integrate the gradient function. Knill’s (2001) analysis seems particularly implausible as a model of human perception given the complexity of its required computations. First, the contours in an image must be fit by a function that is parameterized in terms of arc length. This function must be of a high order with smoothly varying first and second derivatives, so that it is possible to compute the normal and curvature functions along the contour. These must also be parameterized in terms of arc length. In order to compute the local surface normals from these functions, it is then necessary to solve the non-linear differential equation described in Equation 2 for an appropriate set of initial conditions.

Another important property of our planar cut analysis is that it imposes no restrictions on the underlying surface geometry. This stands in stark contrast to the methods proposed by Knill (1992, 2001) and Stevens (1981, 1986), which are only applicable to developable surfaces—i.e., those for which one of the principal curvatures at every point is zero. The evidence is quite clear, however, that images of non-developable surfaces with planar cut contours produce compelling perceptions of 3D shape (Bocheva, 2009; Todd & Oomes, 2002; Todd et al., 2004; Todd & Reichel, 1990; Tse, 2002). The analysis described by Equation 3 is the only one proposed to date that could potentially explain how this is possible.

One of the central components of our planar cut analysis is that the contours must be indexed in order to assign a unique value \( v \) of the texture scale to any given point \( (x, y) \) in an image based on the particular contours in its immediate local neighborhood. The discussion thus far has implied that the contour planes must be equally spaced to create a uniform texture scale, but this is clearly too strong an assumption. Indeed, the stimuli employed in Experiment 1 were all generated with contour planes that were unequally spaced. Note in Figure 2, for example, that...
the black stripes appear narrower than those with a lighter gray color. This perception is most likely based on the fact that the 2D widths of the black image contours are consistently smaller than their neighboring gray contours. This suggests that the relative spacing of contours in local regions of image space could be used to scale the relative spacing of the contour planes in 3D space.

A related issue to consider is how the model might cope with planar cut contours when the inclined planes are not all parallel to one another (e.g., see Figure 7). Todd and Reichel (1990) showed that the perception of 3D shape from contours can tolerate considerable amounts of noise in the orientations of planar cuts, and they speculated that this could be achieved by averaging the orientations and spacing of contours within local neighborhoods of an image. However, there are some contour patterns, such as the one shown in right panel of Figure 7, for which local averaging cannot produce a parallel orientation flow. An alternative approach for this type of pattern is to apply separate analyses for each individual contour. For any fixed Cartesian coordinate system in which the $z$-axis is perpendicular to the image plane, the relative depth of each point along a planar cut contour can be determined from the following equation:

$$Z = \tan\sigma(y\cos\tau - x\sin\tau),$$

where $\sigma$ and $\tau$ are the slant and tilt of the inclined plane. The problem with this approach is that the number of free parameters grows with the number of contours to be analyzed, although the number of constraints grows as well because points of intersection between any two contours must have the same depth (see Ecker, Kutulakos, & Jepson, 2007). As a practical matter, we suspect that this is only feasible for regularly shaped contours (as in the right panel of Figure 7) whose slants and/or tilts can be reliably estimated from their optical projections.

**Model limitations**

Because it was derived using orthographic projection, the planar cut analysis described by Equation 3 will produce systematic distortions when applied to images...
produced with polar projection. These distortions are generally quite small, however, unless a surface is observed (or photographed) with a large viewing angle. For example, Figure 8 shows three images of the same surface under orthographic projection (left panel), polar projection with a $15^\circ$ camera angle (middle panel), and polar projection with a $30^\circ$ camera angle (right panel). A planar cut analysis of these images using a fixed set of parameter settings would produce increasing amounts of estimated surface amplitude as perspective is increased. It should also be noted, however, that observers report a similar variation in the apparent amplitudes of these surfaces that is quite consistent with the predictions of the model.

Another important consequence of assuming orthographic projection is that the estimated sign of curvature is mathematically ambiguous. There are several different ways that this ambiguity could potentially be resolved. If there are smooth occlusions in an image, as in Experiment 2, they can provide reliable information about the sign of curvature in their immediate local neighborhoods (Koenderink, 1984). In the absence of smooth occlusions, the perceived sign of curvature can be influenced by a strong perceptual bias to interpret the pattern of relief so that apparent depth increases with height in the image plane (Langer & Bülthoff, 2001; Mamassian & Landy, 2001; Reichel & Todd, 1990). This is statistically justified by the fact that surfaces are viewed from above more often than they are viewed from below. It is because of this bias that the images in Figure 2 are perceptually stable even though the depicted signs of curvature are mathematically ambiguous. Observers also have a strong bias to perceive surfaces as convex rather than concave (Hill & Bruce, 1994; Langer & Bülthoff, 2001; Liu & Todd, 2004), which is statistically justified by the fact that convex surface patches are more common than concave surface patches. It is interesting to note that when local occlusion information conflicts with these biases it can result in a
perception that is globally inconsistent, much like an impossible figure (Reichel & Todd, 1990).

There are some examples of perceived shape from contour textures that have been published in the literature that are clearly inconsistent with a planar cut analysis. The image in Figure 9 shows a pattern of geodesic contours on an elliptical cylinder (see also Knill, 2001; Todd & Oomes, 2002). Many observers report that the appearance of this image is multistable: It can be seen as a wavy surface with both positive and negative curvatures, or it can appear as a cylinder with a single sign of curvature. The first of these interpretations is consistent with a planar cut analysis, but the second is not. How might this perception of a cylinder be explained? One possibility is that it is based on an analysis such as that of Knill that is designed explicitly to be used with geodesic contours. An alternative possibility is that the relative depths are computed using a planar cut analysis but that the ambiguous signs of curvature are resolved in a piecewise manner, perhaps to conform with the bias for global convexity.

Figure 10 shows another type of contour pattern that produces a strong perception of 3D structure but cannot be analyzed successfully using a planar cut analysis. It is important to keep in mind that when the slants of the inclined planes are close to \( \pm 90^\circ \), Equation 3 cannot be evaluated. The image contours in that case are reduced to a pattern of parallel straight lines that provide no useful information about 3D shape. This is demonstrated quite clearly in the left panel of Figure 10, which shows a developable surface under orthographic projection with parallel planar cuts at a \( 90^\circ \) slant. However, if this same surface is depicted with a sufficient amount of perspective, as shown in the right panel, then its 3D structure is perceived quite clearly (see also Li & Zaidi, 2000, 2004; Todd et al., 2004, 2007). The key source of information in this pattern that is responsible for observers’ perceptions is that the contours converge toward a common vanishing point. This makes it possible to estimate the 3D structure using classical analyses of linear perspective (e.g., Saunders & Backus, 2006), normalized perspective scaling gradients (Gårding, 1992; Malik & Rosenholtz, 1997), or, more likely, by the scaling contrast model recently proposed by Todd et al. (2007). These analyses are primarily effective for planar or asymptotically planar surfaces viewed with relatively large visual angles (see Todd et al., 2007 for a more detailed discussion of this point), whereas the planar cut model is primarily effective for curved surfaces viewed with relatively small visual angles. Thus, these two types of analysis serve as useful complements to one another.

Acknowledgments

This research was supported by grants from the National Science Foundation (BCS-0546107 and BCS-0962119).

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


