The Perception of 3D Shape from Surface Contours

THESIS

Presented in Partial Fulfillment of the Requirements for the Degree Master of Arts in the Graduate School of The Ohio State University

By

Eric James Landon Egan

Graduate Program in Psychology

The Ohio State University

2010

Master's Examination Committee:

Delwin Lindsey

Alex Petrov

James Todd, Advisor
Abstract

A new computational analysis is described that is designed to estimate the shape of an observed surface from the optical projections of its contours in a 2D image. This model assumes that contours on a surface are generated by a series of parallel planar cuts, and it estimates the relative depth between any two surface points based on the number of contour planes by which they are separated, and the apparent 3D orientations of those planes. A psychophysical experiment was performed in an effort to compare the model predictions with the perceptual judgments of human observers about the 3D relief of depicted surfaces. Stimuli consisted of sinusoidally corrugated surfaces with contours that were oriented in different directions. The results reveal that observers’ perceptions vary systematically as a function of contour orientation, and that similar distortions can also be generated by the model with appropriate parameter settings.
Dedicated to…

Sara.

Thank you for moving to Ohio.
Acknowledgments

This Master’s Thesis would not have been possible without the intellectual guidance and support of my advisor, James Todd.

The expertise of Flip Phillips was the integral component required to implement Knill’s model. All of his contributions are greatly appreciated.
Vita

2004.................................................................Moses Brown School

2008.................................................................B.A. Neuroscience, Skidmore College

2008 to present ..............................................Graduate Teaching Associate,

The Department of Psychology,

The Ohio State University

Publications


Fields of Study

Major Field: Psychology
# Table of Contents

Abstract ............................................................................................................................................. ii

Acknowledgments.................................................................................................................................. iv

Vita....................................................................................................................................................... v

List of Tables ........................................................................................................................................ viii

List of Figures ...................................................................................................................................... ix

Chapter 1: Introduction ....................................................................................................................... 1

Chapter 2: Methods .............................................................................................................................. 7

Subjects ............................................................................................................................................... 7

Apparatus .......................................................................................................................................... 7

Stimuli ............................................................................................................................................... 7

Procedure ......................................................................................................................................... 8

Chapter 3: Results ............................................................................................................................. 10

Behavioral Data ................................................................................................................................. 10

Model Implementation ...................................................................................................................... 11

Contours as Lines of Curvature ......................................................................................................... 12
Contours as Geodesics ........................................................................................................... 16

Contours as Planar Cuts ........................................................................................................ 19

Chapter 3: Discussion ............................................................................................................ 23

References ............................................................................................................................. 29
List of Tables

Table 1. The coefficient of determination (R²) and the slope coefficient of the linear regression equation for the versions of Stevens’ model that best fit the ground truth and the average observer depth profiles. ................................................................. 13

Table 2. The R² and the slope coefficient of the linear regression equation for the versions of Knill’s model that best fit the ground truth and the average observer depth profiles. . 17

Table 3. The R² and the slope coefficient of the linear regression equation for the best fitting versions of planar cut model to both the ground truth and the average observer. . 20
List of Figures

Figure 1. A pattern of contours that is perceived as a smoothly curved 3D surface. .......... 1

Figure 2. Measurable properties of an image that can be used to predict the relative depth between two points (red circles) on a surface rendered with a planar cut contour texture. The axes are aligned with the approximated tilt ($\tau$) of the contour planes that generated the texture. $D_y$ is the distance in the direction of the y-axis that separates the two points. $C_x$ is the number of contour planes that separate the points along the x-axis (i.e. 1), and $C_y$ is the number of contour planes that separate them along the y-axis (i.e. -6). The slant ($\sigma$) and separation ($S$) of the contour planes must also be approximated. ......................... 4

Figure 3. Example images from the five stimulus conditions used in the experiment. One 3D surface was used to render all images; only the orientation of the planar cut contour texture was manipulated. ........................................................................................................... 5

Figure 4. A screen capture for an example trial of the profile reproduction task. Using a mouse, participants adjusted the heights of the circles on the left monitor to match the perceived relative depths of the corresponding circles on the right monitor. ..................... 9

Figure 5. The five graphs show the collapsed data for all observers in each of the five stimulus conditions (A-E). The red circles represent the average adjustments made by all observers for the profile reproduction task. The black contour represents the ground truth. The x- and y-axes of the graphs denote the x-axis and the ambiguous z-axis, respectively, of each stimulus. ........................................................................................................... 11
Figure 6. The two graphs show the depth profiles produced by each of the three models that best fit the observers’ averaged data for stimulus conditions B and D........................ 14

Figure 7 The black lines are the planar cut model’s parameters that best fit the behavioral data for each of the 5 conditions with error bars denoting one standard error. The red (short dashed) line marks the ground truth. The blue (long dashed) line marks the linear orientation of the contours in the image. ........................................................................................................ 21

Figure 8. Both images are of the same surface with a geodesic contour texture viewed with negligible perspective. Both images have four perceptually possible depth profiles shown down the center. The different profiles are more easily perceived for the image on the right since it has no smooth occlusion boundaries....................................................... 25

Figure 9. Both images are of the same cylindrical surface with a planar cut contour texture. The image on the left was rendered with negligible perspective while the image on the right was rendered with a large amount of perspective. .......................................................... 27
Chapter 1: Introduction

One of the most perceptually compelling methods for pictorial depiction of smoothly curved surfaces involves regular patterns of surface contours. Consider, for example, the image presented in Figure 1. At one level of analysis, this image consists of nothing more than a pattern of wavy lines in the picture plane, but it appears perceptually as a smoothly curved surface in three-dimensional space. The depiction of curved surfaces using patterns of image contours is especially common in engineering drawing and in optical art, and it has also been the subject of several psychophysical investigations (Li & Zaidi, 2000; Mamassian & Landy, 2001; Reichel & Todd, 1990; Stevens & Brooks, 1987; Todd & Reichel, 1989, 1990; Todd & Oomes, 2002; Todd, Thaler & Dijkstra, 2005).

Figure 1. A pattern of contours that is perceived as a smoothly curved 3D surface.
How is it possible to determine the shape of a surface from a pattern of image contours? The first formal analysis of this problem was presented by Stevens (1981, 1986), who considered the special case of developable surfaces whose contours are aligned with directions of principal curvature. Whenever a surface conforms to these restrictions, it is possible to compute the overall pattern of local surface orientations up to a one parameter family of possible interpretations. A similar analysis was later developed by Knill (1992, 2001), which loosened these restrictions somewhat by allowing the contours on a developable surface to be geodesics rather than lines of curvature. Under those conditions, it is possible to compute the overall pattern of surface orientations within two degrees of freedom, which specify the global orientation of a surface. Because all lines of curvature are also geodesics, Stevens’ analysis constitutes a special case of Knill’s, in which the set of possible global orientations is more constrained.

A third possible approach to this problem has more recently been suggested by Tse (2002). He argued that observers’ perceptions of shape from contour textures are based on an implicit assumption that the contours form parallel planar cuts through a surface. Tse provided numerous demonstrations to show that the perception of shape from contours is not limited to developable surfaces as suggested by the analyses of Stevens (1981, 1986) and Knill (1992, 2001). However, he did not present a specific computational model with which the 3D structure of a surface can be determined from its pattern of projected contours.
In an effort to expand on Tse’s hypothesis, it is useful begin by analyzing the special case of a surface that is textured with a series of equally spaced parallel planar cuts that are all perpendicular to the observer’s line of sight (e.g. see Figure 1). The contours in that case can be used as a ruler because the difference in depth ($D_z$) between any two surface points will always be proportional to the number of contour planes that separate them. Thus, for any given point pair:

$$D_z = CS$$  \hspace{1cm} (1)$$

where $C$ is the number of contour planes by which they are separated, and $S$ is a scaling factor that is defined by the physical separation between adjacent contour planes.

Let us now consider the more general case of a surface whose contour planes are all oriented at a slant $\sigma$ and a tilt $\tau$. For convenience, we will align the $y$-axis of the image plane coordinate system to be parallel to the tilt of the contour planes. For images rendered with negligible perspective, the difference in depth between any two points can be computed from the following equation:

$$D_z = \frac{D_y \sin \sigma + C_x S + C_y S}{\cos \sigma}$$  \hspace{1cm} (2)$$

where $D_y$ is the difference between their $y$-coordinates, $C_x$ is the number of contour planes that separate the points along the $x$-axis, and $C_y$ is the number of contour planes that separate them along the $y$-axis (see Figure 2). It is important to recognize that $D_y$, $C_x$ and $C_y$ are all measurable properties of an image, whereas $\sigma$, $\tau$ 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 ±90°, then Equation 2 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.

![Diagram of measurable properties](image)

**Figure 2.** Measurable properties of an image that can be used to predict the relative depth between two points (red circles) on a surface rendered with a planar cut contour texture. The axes are aligned with the approximated tilt ($\tau$) of the contour planes that generated the texture. $D_y$ is the distance in the direction of the y-axis that separates the two points. $C_x$ is the number of contour planes that separate the points along the x-axis (i.e. -1), and $C_y$ is the number of contour planes that separate them along the y-axis (i.e. 6). The slant ($\sigma$) and separation ($S$) of the contour planes must also be approximated.

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). How then do 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.

![Figure 3. Example images from the five stimulus conditions used in the experiment. One 3D surface was used to render all images; only the orientation of the planar cut contour texture was manipulated.](image)

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 3 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 the present article 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.
Chapter 2: Methods

Subjects

Seven observers participated in the experiment. They were all naïve to the issues being investigated and all had normal or corrected to normal vision.

Apparatus

The experiment was conducted on a Dell Dimension 8300 PC with an ATI Radeon 9800 PRO Graphics card. Stimuli were presented on two monitors, a 19 inch CRT and a 19 inch LCD monitor. Both monitors had a spatial resolution of 1280 X 1024 pixels. The stimulus images were presented at native resolution within a 25.4 X 19.0 cm region (800 X 600 pixels) on the right (LCD) display screen, which subtended 14.6° X 10.9° of visual angle when viewed at a distance of 100 centimeters. Observers wore an eye patch for monocular viewing.

Stimuli

Images of a sinusoidally corrugated surface were rendered using 3D Studio Max by Autodesk with procedural parallel 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° and a global tilt of 25°.

The textures were created by a volume of alternating light and dark rectangular slabs which produced a series equally spaced parallel planar cuts through the surface. The
planar cuts had slants and tilts of (-62°, -54°), (-53°, -37°), (-49°, -19°), (-47°, 0°) and (-49°, 19°), respectively, for each of the five stimulus categories. 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 3. Stimulus condition B is a special case where the contours were lines of maximum curvature.

**Procedure**

A profile reproduction task was employed in order to measure perceived relative depth between multiple surface locations. An example screen capture of what a trial might look like is shown in Figure 4. On each trial a stimulus was presented 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. A handheld 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. These five scan lines were evenly distributed so that each consecutive pair of scan lines was separated by 100 vertical image pixels. A random presentation order was used to ensure each stimulus condition received responses from each scan line. 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 sinusoidal surface. The
dots of each scan line corresponded to the same six phase intervals with respect to the sinusoidal 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.

![Figure 4](image.png)

Figure 4. A screen capture for an example trial of the profile reproduction task. Using a mouse, participants adjusted the heights of the circles on the left monitor to match the perceived relative depths of the corresponding circles on the right monitor.

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. The experiment began when both parties were confident that the observer understood the task. Each observer was presented with all twenty-five stimuli: five variations of each of the five stimulus categories. The mean time per trial was between one and two minutes. All observers completed the experiment within a thirty to forty-five minute time period.
Chapter 3: Results

*Behavioral Data*

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 different scan lines into a single apparent profile in depth. An analysis of linear regression revealed that the average correlation between the profiles generated by each pair of observers was quite high with a coefficient of determination ($R^2$) of .79. Thus, all subsequent analyses and model fits were performed on depth profiles that were computed from the average judgments of all observers.

The results obtained for each of the different contour orientations are presented in Figure 5, which shows the average response profile in each condition (red circles) together with the ground truth (black curves). Error bar are not shown, but the radius of each circle is the size the largest standard error found for all conditions. Note in this figure that the response curves are qualitatively similar to the ground truth, but that there are two characteristic patterns of distortion. First, the overall magnitude of apparent depth is systematically underestimated, which is a common finding in many experiments on perceived surface shape (Todd, Oomes, Koenderink, & Kappers, 2001). Second, in many of the conditions there is a systematic shearing of the judged depth profiles relative to the
ground truth (Koenderink, van Doorn, Kappers, & Todd, 2001). For example, the mean response for Condition A was sheared, such that the left side of the surface appeared closer in depth than the right side. The opposite is true for Condition E where the right side of the surface appeared closer in depth than the left side.

**Model Implementation**

Additional analyses were performed to determine whether these systematic distortions in the behavioral data could be explained by any of the three models described in the introduction. This was achieved by implementing each model for the five different contour patterns employed in the present experiment (see Figure 3), and computing the values of their free parameters that provide the best least-squares fits to the average response profiles.

**Figure 5.** The five graphs show the collapsed data for all observers in each of the five stimulus conditions (A-E). The red circles represent the average adjustments made by all observers for the profile reproduction task. The black contour represents the ground truth. The x- and y-axes of the graphs denote the x-axis and the ambiguous z-axis, respectively, of each stimulus. The radius of the circles denotes the greatest standard error.
**Contours as Lines of Curvature**

The first computational model for estimating 3D shape from contour textures was proposed by Stevens’ (1981, 1986). His analysis was based on three critical assumptions: First, the underlying surface must be developable; second, the surface contours must be lines of maximum curvature; and third, the images to be analyzed must be 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:

\[
N = \{-\alpha \sin\beta, \cos\beta - \frac{\alpha^2 + 1}{\alpha}, \sin\beta\}
\]

(3)

where \( \beta \) is the 2D angle between the 2D ruling vector and the vector tangent to the 2D contour and \( \alpha \) is the \( z \)-component of the surface ruling vector (equivalent to the direction of zero curvature). It is important to note that \( \alpha \) is the only free parameter in this equation.

To allow for easier comparisons with other models and our data, the model was altered to produce depth profiles phase aligned with our behavior data. This manipulation involves taking a surface normal function obtained from Equation 3 that represents a horizontal scan line and extracting a depth gradient. The depth gradient can then be integrated to produce a depth profile. An analysis was performed using Stevens’ model to determine the value of \( \alpha \) in each condition that provided depth profiles with the best least-squares fits to the ground truth. The depth profiles were calculated for all values of \( \alpha \) within half a degree of accuracy. A linear regression was then performed to compare the best fitting profiles to the ground truth.
Let us first consider how this model performs when given an image from Condition B, which is the only condition that satisfies the model’s underlying assumptions. The results revealed that the model shows little variation from the ground truth with an $R^2 > .99$. Besides this correlation, it is also important to consider the slope coefficient of the linear regression equation. This slope represents the average ratio of the predicted depths over the ground truth depths. A value of one is an accurate slope, while values above and below one are over and underestimations of global depth, respectively. For Condition B, the slope was 1.01, indicating the model is accurate in terms of scaling in depth. The minor differences between the model’s results and the ground truth are caused by the limited resolution of our implementation. This test confirms that Stevens’ model can accurately calculate 3D shape from contours for a surface that conforms to its underlying assumptions.

Table 1. The coefficient of determination ($R^2$) and the slope coefficient of the linear regression equation for the versions of Stevens’ model that best fit the ground truth and the average observer depth profiles.

<table>
<thead>
<tr>
<th>Condition</th>
<th>R-Squared</th>
<th>Slope Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ground Truth</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>.93</td>
<td>3.24</td>
</tr>
<tr>
<td>B</td>
<td>.99</td>
<td>1.01</td>
</tr>
<tr>
<td>C</td>
<td>.99</td>
<td>1.46</td>
</tr>
<tr>
<td>D</td>
<td>.99</td>
<td>11.5</td>
</tr>
<tr>
<td>E</td>
<td>.58</td>
<td>&gt;&gt;100</td>
</tr>
<tr>
<td>Average Observer</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>.96</td>
<td>2.26</td>
</tr>
<tr>
<td>B</td>
<td>.97</td>
<td>1.85</td>
</tr>
<tr>
<td>C</td>
<td>.98</td>
<td>3.04</td>
</tr>
<tr>
<td>D</td>
<td>.96</td>
<td>6.60</td>
</tr>
<tr>
<td>E</td>
<td>.86</td>
<td>&gt;&gt;100</td>
</tr>
</tbody>
</table>
However, the model is not robust to violations of these assumptions. The top of Table 1 shows the $R^2$ and the slope coefficient of the linear regression equation for the depth profiles that best fit the ground truth (for all possible values of $\alpha$ with half a degree of accuracy). For the conditions in which the contours were not lines of curvature, the correlations were high with $R^2 > .93$ for all conditions except Condition E where the model could not account for more than 58% of the variance. For the remaining conditions (A, C & D), the slopes of the regression equations are greater than one. As the contours are rotated farther from following lines of maximum curvature the model simulations increasingly overestimate the global depth of the surface. Remember, one of the systematic distortions found in our behavioral results was an underestimation of global depth but the model is producing overestimated depth profiles.

Figure 6. The two graphs show the depth profiles produced by each of the three models that best fit the observers’ averaged data for stimulus conditions B and D.
One final analysis was performed using Stevens’ model to determine the value of $\alpha$ (with half a degree of accuracy) in each condition that provided the the depth profiles with the best least-squares fits to the average observer depth profiles. The depth differences for each of the six phase intervals in each condition from the best fitting profiles were compared to the average judged depth differences using an analysis of linear regression. The results shown in the bottom of Table 1 reveal that Stevens’ model can account for a large amount of the variance among the judged depth differences between adjacent probe points with a mean $R^2 > .94$. These high correlations are only accurate up to scaling in depth, with regression line slopes starting at 1.86 for Condition B and growing into triple digits for Condition E.

For a graphical representation of the simulations, the black (long dashed) contours in Figure 6 show the computed depth profiles for Conditions B and D. Note that for Condition B (assumptions met) compared to Condition D (assumptions violated), the contour is much closer to the average observer’s data as shown by the red circles. The global depth for Condition D is overestimated by a factor of 6.60. It is evident from the poor estimations of global depth that this model cannot explain the specific patterns of perceptual distortion that are evident in Figure 5.
Contours as Geodesics

Let us now consider the model proposed by Knill (1992, 2001). Like Stevens’ model, Knill’s analysis assumes the underlying surface is developable, and that it is viewed with negligible perspective. However, it is based on a more general 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:

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

where \(k(s)\) 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. The notation \(\langle x, y \rangle\) in this equation signifies the inner product between two vectors, and \(x \wedge y\) signifies the outer product.

Although Knill’s analysis is mathematically quite elegant, he provides almost no discussion about how the model might be implemented in practice to compute 3D shape from actual image data, as opposed to analytically defined surfaces. There are numerous problems in implementing this approach that limit its practical utility, and make it highly implausible as a biological model. 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. Finally, the nonlinear differential equation described in Equation 4 must be solved for a given set of initial conditions, which requires a known (or estimated) surface normal at one point along the contour.

Table 2. The $R^2$ and the slope coefficient of the linear regression equation for the versions of Knill's model that best fit the ground truth and the average observer depth profiles.

<table>
<thead>
<tr>
<th>Condition</th>
<th>R-Squared</th>
<th>Slope Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>.99</td>
<td>.884</td>
</tr>
<tr>
<td>B</td>
<td>.99</td>
<td>1.01</td>
</tr>
<tr>
<td>C</td>
<td>.99</td>
<td>.930</td>
</tr>
<tr>
<td>D</td>
<td>.99</td>
<td>.891</td>
</tr>
<tr>
<td>E</td>
<td>.99</td>
<td>.884</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ground Truth</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>D</td>
</tr>
<tr>
<td>E</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Average Observer</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
</tr>
<tr>
<td>B</td>
</tr>
<tr>
<td>C</td>
</tr>
<tr>
<td>D</td>
</tr>
<tr>
<td>E</td>
</tr>
</tbody>
</table>

Similarly to Stevens’ model, the surface normal function obtained from Equation 4 for a scan line can be used to generate a depth gradient that can then be integrated to compute a depth profile. As mentioned above, the model’s predictions vary depending on the 3D surface normal, with two degrees of freedom, given as an initial condition for Equation 4. An analysis was performed to determine what combinations of the two free parameters (within half a degree of accuracy) in Knill’s model, produced depth profiles for each condition that best fit the ground truth. A linear regression was performed to compare the best fitting profiles to the ground truth. The results, shown in the top of Table 2, revealed that the model is capable to some extent of closely approximating the
ground truth. Condition B was the only condition with an accurate regression line slope and this is not surprising considering it is the only condition where the contours are actually geodesics.

It should be noted that, for developable surface like our stimuli, geodesics and planar cut contours are quite similar (Knill, 2001), but this is not the case for most surfaces. The other four predictions likely suffered due to the differences between planar cuts and geodesics as well as the sensitive nature of high order functions used by the model. Still, Knill’s model is capable of producing depth profiles similar to the ground truth up to a scaling in depth for all five conditions.

Let us now consider how well Knill’s model can capture the specific patterns of perceptual distortion found in our behavioral results. In a similar fashion to Stevens’ model, an analysis was also performed to determine what combinations of the two free parameters (with half a degree of accuracy) in Knill’s model, produced depth profiles that best fit average observers’ depth profiles. Again, the depth differences for each of the six phase intervals in each condition from the best fitting profiles were compared to the average judged depth differences using an analysis of linear regression.

The results of the linear regression are shown in the bottom of Table 2. Knill’s model produces high correlations for Conditions B, C & D, and lower performances for the remaining conditions. These correlations are only accurate up to scaling in depth, which at best, was off by factor of 1.27 in Condition B, and at worst by a factor greater than 2 in Condition A. The blue (medium dashed) contours in Figure 6 show the computed depth profiles for Conditions B and D.
Contours as Planar Cuts

The final model we will consider is the one described in Equation 2. It is based on the assumption that the contours on a surface are formed by a series of parallel planar cuts (see also Tse, 2002). This model has three free parameters: \( \sigma \), \( \tau \) and \( S \), which correspond to the slant, tilt and physical separation of the contour planes. When given the experimental stimuli along with correct values of the parameters the model produced accurate reconstructions of the ground truth according to the results of a regression analysis shown in the top of Table 3. There was almost zero variation between the model’s prediction and the ground truth with a mean \( R^2 > .99 \) and perfect depth scaling with a mean regression slope of 1.00. This result was consistent with our expectations, as none of the model’s underlying assumptions were violated. Any slight variations compared to the ground truth are due to the limited resolution of our implementation.

The stimuli from the present experiment were analyzed using this model to determine the parameter values of \( \sigma \), \( \tau \) and \( S \) that provided the depth profiles with the best least-squares fits to the average observer profiles. The parameter search included all values of \( \sigma \) & \( \tau \) to within 0.1° of accuracy and values of \( S \) from 0 to 2 in steps of 0.05. The depth differences for each of the six phase intervals in each condition from the best fitting profiles were compared to the average observers’ judged depth differences using an analysis of linear regression. The results shown in the bottom of Table 3 revealed that this model accounted for on average 98% of the variance among the judged depth differences between adjacent probe points for the five conditions. The model slightly...
underestimated the overall depth perceived with a mean linear regression equation slope of .977. The green (short dashed) contours in Figure 6 show the computed depth profiles for Conditions B and D. While the planar cut model slightly underestimated the global depth of the profiles, these predictions are the closest to the average observer’s data when compared to the other models implemented.

Table 3. The $R^2$ and the slope coefficient of the linear regression equation for the best fitting versions of planar cut model to both the ground truth and the average observer.

<table>
<thead>
<tr>
<th>Condition</th>
<th>R-Squared</th>
<th>Slope Coefficient</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ground Truth</td>
<td>A</td>
<td>.99</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>.99</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>.99</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>.99</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>.99</td>
</tr>
<tr>
<td>Average Observer</td>
<td>A</td>
<td>.98</td>
</tr>
<tr>
<td></td>
<td>B</td>
<td>.99</td>
</tr>
<tr>
<td></td>
<td>C</td>
<td>.99</td>
</tr>
<tr>
<td></td>
<td>D</td>
<td>.98</td>
</tr>
<tr>
<td></td>
<td>E</td>
<td>.96</td>
</tr>
</tbody>
</table>
An additional analysis determined the parameter values of $S$, $\sigma$ and $\tau$, for each individual observer that produce the best least-squares fits to their judged depth profiles. When a linear regression was performed between the best fits and each individual’s data, the results showed that the model accounts for on average 94% of the variance in each individual’s depth judgments. The top, middle and bottom of Figure 7 shows how the best fitting values of $S$, $\sigma$ and $\tau$, respectively, varied across the different experimental conditions for all observers relative to the ground truth. The error bars denote one standard error between observers for a given parameter. The largest deviations from the ground truth occurred for the scaling parameter $S$, which required a value of one in order to produce veridical performance. Instead, the values obtained in the different conditions had a mean of 0.51, which reflects the overall underestimation of

![Figure 7](image.png)

Figure 7: The black lines are the planar cut model’s parameters that best fit the behavioral data for each of the 5 conditions with error bars denoting one standard error. The red (short dashed) line marks the ground truth. The blue (long dashed) line marks the linear orientation of the contours in the image.
surface relief in the observers’ perceptions. Incorrect values of $\sigma$ and $\tau$ produce shearing distortions in the computed depth profiles. The best fitting values of $\sigma$ were reasonably accurate for conditions A, B and C, and slightly underestimated for Conditions D and E. The best fitting values of $\tau$, in contrast, varied much less among the different conditions than did the actual values used to generate the stimuli. One possible explanation of this is that $\tau$ was approximated by the overall orientations of the contours in the image plane. To test this possibility, an analysis of linear regression was performed on the contour coordinates in the image plane to determine the overall contour orientation for each condition. The orientations revealed by this analysis are indicated by the blue (long dashed) contour in the bottom of Figure 7. Note that this provides a reasonably good prediction of the best fitting values of $\tau$, except that they are slightly overestimated for Conditions A and B.
Chapter 3: Discussion

The research described in the present article considers observers’ depth judgments of images depicting surfaces with planar cut contour textures. The results revealed that these judgments are stable across multiple observers, systematically underestimated and sometimes sheared relative to the ground truth.

We examined three models of estimating 3D shape from contour textures to see how well they can capture human perceptual abilities. In order to assess these models, we will judge them on several criteria. First, how well do the models fit the behavioral data? Second, how many free parameters do the models have? Third, what order measurements do the models require? Fourth, are the models computationally plausible from a biological standpoint? Finally, how general or restricted are the models?

A model’s ability to fit behavioral data is one of the most important factors when comparing models. Our results show that all of the models can approximate the general shape and shearing of observers’ judgments but differ with respect to the overall magnitude of global depth relief. The scaling in depth for Stevens’ model is off by a factor of five or more for some images. Knill’s model performs better, but remains off by a factor of at least 1.27 and up to a factor of 2.12. The planar cut model is the best fitting model with all fits accurate within a 5% scaling in depth.
Stevens’ model has only a single free parameter so it is not surprising that fits are not optimal. Knill’s model, with a second free parameter, has much better fits compared to Stevens’ model. The planar cut model has the most free parameters, however one of these parameters is likely estimated by the orientation of the contours in the image plane.

The order of the measures required by a model is important when judging its biological plausibility. Higher order measurements are more difficult and are more susceptible to noise. Our model involves only zero order measurements of distance in the image plane and distance in the texture space. Stevens’ model uses first order measurements including the orientation of the surface ruling and local orientation of a contour. Knill’s model is the most complex requiring the first order measurements of orientation and second order measurements of local curvature at each point along a contour.

Besides the measurements required, the biological plausibility of the computation involved in an analysis should also be considered. Our model computes depth directly from a relatively simple algebraic equation. Stevens’ model computes local orientation rather than depth, but depth can be obtained by extracting the depth gradient from the local surface normals and then integrating the gradient function. Knill’s model is by far the most complex, requiring contours to be fit with a high order function that allows for smoothly varying first and second derivatives that are defined in terms of arc length. These second derivatives functions are then used to solve the differential equation shown in Equation 4 to compute local surface orientation. Finally, like Stevens’ model, a depth gradient must be extracted and integrated to obtain a pattern of surface depth.
Lastly, the models must be discussed in terms of generalizability. Stevens’ and Knill’s models are highly restricted since they require to developable surfaces. Alternatively, our model is generalizable for any surface.

![Figure 8](image.png)

Figure 8. Both images are of the same surface with a geodesic contour texture viewed with negligible perspective. Both images have four perceptually possible depth profiles shown down the center. The different profiles are more easily perceived for the image on the right since it has no smooth occlusion boundaries.

To summarize, although our model has more free parameters, it is clearly better by all other criteria when compared to the alternatives. It has shown the best in terms of correlation and scaling in depth while using only zero order image measures in a relatively simple algebraic equation. Furthermore, it is not restricted to certain surface types. With that being said, there are two counter cases that deserve consideration where the planar cut model cannot account for human perception.
The first counter case that our model is not specifically designed to handle involves surfaces with geodesic contours. While earlier we mentioned that planar cuts closely resemble geodesics for our stimuli this is not the case for most surfaces. The left side of Figure 8 is a cylindrical surface with geodesic contours viewed with negligible perspective that is clearly perceivable as 3D. If the same surface was textured with planar cut contours, the contours would have zero curvature in the image plane and the surface would appear as a series of flat parallel lines. The geodesic paths of the contours in the image do produce contour curvature in the image plane and the surface is often perceived as a convex cylinder. The sign of relief is actually ambiguous and there are four possible depth relief interpretations as shown down the center of Figure 8. Koenderink, & van Doorn, (1982) suggest the smooth occlusion boundaries on either side of the cylinder create a bias to perceive the surface as convex. The effects of the smooth occlusion boundaries are diminished when backdrop texture is removed and the contours are more widely spaced like those in the right side of Figure 8. These contours are identical in curvature to those on the left but the surface is often perceived as multi-stable. It can still be perceived as a convex cylinder or more commonly as flowing ribbon. This second interpretation matches the depth relief produced by the planar cut model. However, the model can also account for the convex (and concave) perceptions of the surface if it is applied to the image in a piecemeal manner where the ambiguous sign of relief is reversed for one side.

It should be noted that stimuli used in the current experiment are ambiguous up to a sign of relief, and so are the predictions of all three models discussed. Todd & Reichel
(1990) suggest the visual system often disambiguates such images by interpreting the overall surface slant so that depth increases with height in the image plane. This is otherwise known as a ground plane bias.

Figure 9. Both images are of the same cylindrical surface with a planar cut contour texture. The image on the left was rendered with negligible perspective while the image on the right was rendered with a large amount of perspective.

Finally, our model cannot be evaluated when the contour planes are close to being perpendicular with the image plane for images with negligible perspective. However, when this is the case, humans will not perceive any depth either. The identical surface used to create Figure 8 was used to create the left side of Figure 9. The contour planes are perpendicular to the image plane, and still viewed with negligible perspective. Interestingly, when the same scene is shown with large amounts of perspective, as in the right side of Figure 9, the surface is perceived as 3D. The planar cut model presented here cannot account for surfaces viewed with perspective. Todd and colleagues (2007)
proposed a model based on scaling contrast that accounts for human perceptions of textured surfaces viewed with perspective.
References


