Qualitative perception of 3D shape from patterns of luminance curvature

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

Ying Yu Department of Psychology, University of Giessen, Giessen, Germany

Flip Phillips School of Film and Animation, Rochester Institute of Technology, Rochester, NY, USA

A new source of information is proposed for the perception of three-dimensional (3D) shape from shading that identifies surface concavities from the curvature of the luminance field. Two experiments measured the abilities of human observers to identify concavities on smoothly curved shaded surfaces depicted with several different patterns of illumination and several different material properties. Observers were required to identify any apparent concavities along designated cross sections of the depicted objects and to mark each concavity with an adjustable dot. To analyze the results, we computed both the surface curvature and the luminance curvature along each image cross section. The results revealed that most responses were in concave regions of the luminance profiles, although they were often shifted in phase relative to the curvature of the depicted surfaces. This pattern of performance was surprisingly robust over large changes in the pattern of illumination or surface material properties. Our analysis predicts that observers should make false alarm responses in regions where a luminance concavity does not correspond to a surface concavity, and our empirical results confirm that prediction.

Qualitative perception of three-dimensional shape from patterns of luminance curvature

Human observers have a remarkable ability to perceive the three-dimensional (3D) shapes of objects from patterns of image shading (i.e., chiaroscuro), and manipulations of these patterns have been used to create the appearance of 3D shape in pictorial art for over two millennia. Figure 1 shows a pebble mosaic that was discovered by archeologists in the northern Greek city of Pella, which was the capital of Macedonia during the reign of Alexander the Great in the fourth century BCE. Note how the systematic variations in shading provide a sense of volume to the depicted surfaces. This is particularly clear in the depiction of the cloth capes.

Although observers' perceptions of 3D shape from shading have been investigated for almost five decades, we know surprisingly little about the relevant information that makes this phenomenon possible. The first computational models for determining shape from shading were developed in the 1970s by Bertholdt Horn and his colleagues at MIT, and this quickly became a cottage industry within the field of computer vision (e.g., see Horn, 1975; Ikeuchi & Horn, 1981; Pentland, 1984; Lee & Rosenfeld, 1985). These models were designed to computationally invert the physical process by which light reflects from an object to determine the local orientation of each visible surface region. The problem with this approach is that it requires numerous assumptions that are seldom if ever satisfied in the natural environment. For example, most models assume that surfaces scatter light uniformly in all directions with no specular highlights and that they are illuminated from a single direction with no cast shadows or interreflections from other surfaces. It was later shown by Belhumeur, Kriegman, and Yuille (1999) that shading information under these conditions is inherently ambiguous, such that any given pattern of shading is consistent with an infinite family of possible interpretations that are all related by affine shearing or stretching transformations in depth.

Numerous psychophysical experiments have measured observers' perceptions of 3D shape from shading. Many of these have investigated the stability of perceived shape over changes in the pattern of illumination (e.g., Caniard & Fleming, 2007; Christou & Koenderink, 1997;

Citation: Todd, J. T., Yu, Y., & Phillips, F. (2023). Qualitative perception of 3D shape from patterns of luminance curvature. *Journal of Vision*, *23*(5):10, 1–16, https://doi.org/10.1167/jov.23.5.10.

https://doi.org/10.1167/jov.23.5.10 Received January 4, 2023; published May 12, 2023 ISSN 1534-7362

Figure 1. A pebble mosaic of a stag hunt from 300 BCE by Gnosis. This is one of the earliest known examples of the use of shading (i.e., chiaroscuro) for the depiction of 3D shape in pictorial art.

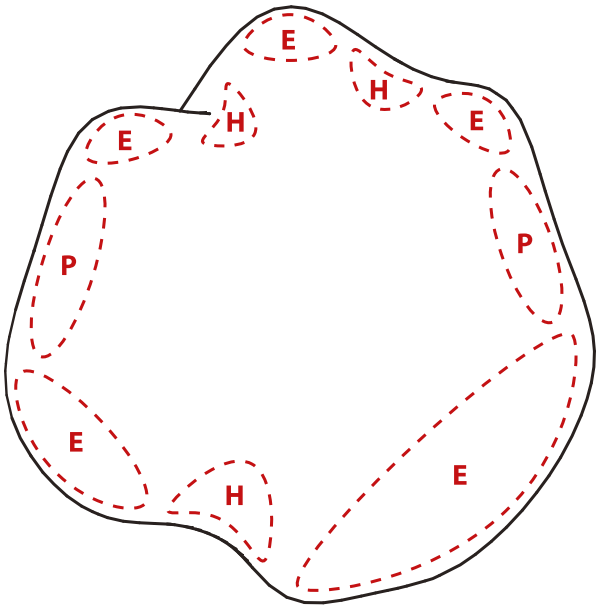

Figure 2. The rim of a smoothly curved object. The regions outlined by dashed lines in the interior are labeled as elliptic (E), hyperbolic (H), and parabolic (P) based on the curvature along the contour.

Curran & Johnston, 1996; Egan & Todd, 2015; Koenderink et al., 1996a, 1996b; Nefs et al., 2005), and the results have confirmed that observers' shape judgments are often distorted by an affine transformation in depth relative to the ground truth. For example, if the primary source of illumination is to the right of an object, the surface regions on the right will appear slightly closer in depth, whereas those on the left will appear slightly farther away. The opposite effect will occur if the object is illuminated from the left. Note that this is perfectly consistent with the bas-relief ambiguity described by Belhumeur, Kriegman, and Yuille (1999).

It is important to recognize that these affine distortions have no effect at all on the overall pattern of surface concavities and convexities. This suggests that a good way of achieving shape constancy for ambiguous stimuli might be to base the appearance of 3D shape on the qualitative aspects of surface curvature that remain relatively stable over changes in viewing conditions (see also Koenderink & van Doorn, 1980; Kunsberg & Zucker, 2018, 2021), rather than a map of local surface depths or orientations, which is much less stable. One justification for this approach is that observers' verbal descriptions of 3D surfaces typically involve the identification of topographic features, such as bumps, dimples, ridges, valleys, and saddles, and how those features are arranged with respect to one another. Note that all of these features are defined by the pattern of surface concavities and convexities.

One source of information about qualitative 3D shape is provided by an object's smooth occlusion contour, which is also referred to as the rim (Koenderink, 1984; Koenderink & van Doorn, 1982). The surface curvature perpendicular to the rim (on the attached side) is always convex. Thus, if a contour bends inward toward the object, then an attached surface region has a positive Gaussian curvature like a sphere (i.e., it is elliptic). If the contour bends outward away from the object, then an attached surface region has a negative Gaussian curvature like a saddle (i.e., it is hyperbolic), and if it is straight, then an attached surface region has zero Gaussian curvature like a cylinder (i.e., it is parabolic). Koenderink and van Doorn (1982) also showed that when the rim ends abruptly in an image, that can only occur in a hyperbolic region. Figure 2 shows the rim of a smoothly curved object presented in isolation. The red dashed curves in this figure show the signs of Gaussian curvature in adjacent regions along the contour. Note that the inner boundaries of these regions cannot be determined based solely on the rim. That requires an analysis of the image shading.

This article will consider a second source of qualitative information about surface curvature that relies on the pattern of shading but can also be used in conjunction with smooth occlusion contours. The upper left panel of Figure 3 shows a uniformly convex surface that is illuminated from the right. The two

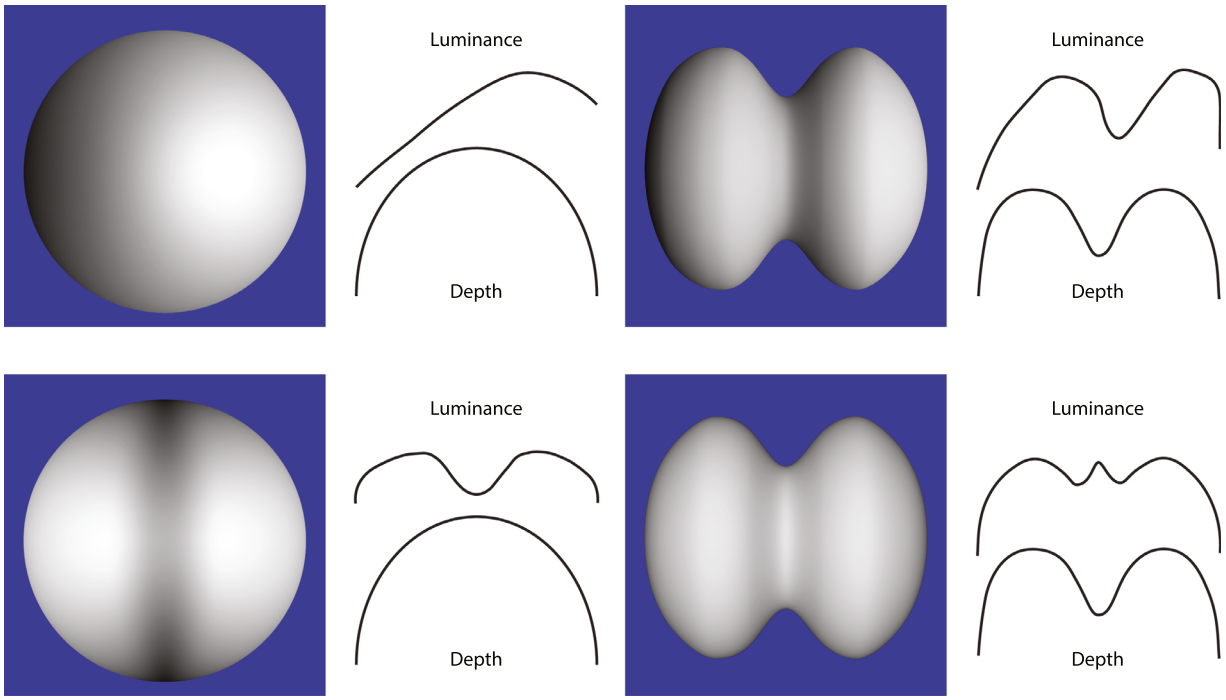

Figure 3. Four surfaces with Lambertian shading. The curves to the right of each object show the depth and luminance profiles of a horizontal cross section through its center. For the curves in the top row, the depth and luminance profiles have the same number of concavities. For the ones in bottom row, the luminance profiles have a greater number of concavities than the depth profile.

curves to the right of the image show the depth and luminance profiles along a horizontal cross section through the center of the object. Note that they are both uniformly convex, but they are not aligned with one another. The peak of the depth profile is at the center of the object, whereas the peak of the luminance curve is shifted toward the direction of illumination. With generic patterns of illumination, a uniformly convex luminance pattern can only occur for uniformly convex surfaces. The fact that the occlusion contour is also uniformly convex provides converging information to support that interpretation.

The upper right panel of Figure 3 shows an hourglass shape illuminated from the right to demonstrate what happens when a surface contains both convex and concave regions. The two curves to the right of the image show the depth and luminance profiles along a horizontal cross section through the center. Note that the two curves have similar shapes but that the phase of the luminance profile is shifted toward the direction of illumination. Along any cross section through a surface, the regions immediately adjacent to a smooth occlusion contour will always be convex, and each concavity will always be bounded on both sides by a convexity. If the depth profile along a surface cross section contains a concavity, then the luminance profile will contain a concavity as well (unless the illumination is specifically tailored to eliminate that).

These observations suggest that the presence of luminance concavities in an image could provide useful information about the presence of surface concavities. Consider the pebble mosaic shown in Figure 1, where all of the depicted concave regions are colored as dark luminance concavities, and all of the depicted convex regions are colored as light luminance convexities. Although this is a surprisingly effective artistic device, the shading observed in the natural environment is much more complex.

Exploiting information from luminance curvature in more natural settings poses a number of difficulties. For example, in order to localize surface concavities, it is necessary to compensate for possible phase shifts between the depth and luminance profiles along any surface cross section. Todd and Reichel (1989) argued that this could be achieved by comparing the luminance at points on the occlusion boundary that are at opposite ends of a surface cross section. Note that the luminance profile in the upper right panel of Figure 3 is much brighter on the right than on the left, which indicates that the illumination is heavily skewed to the right. Thus, the deepest part of the depth profile should be located on the left side wall of the luminance concavity that faces toward the direction of illumination.

Another problem that needs to be addressed for identifying surface concavities from luminance curvature is that luminance concavities can arise for a variety of reasons, and the number of concavities along

a luminance profile is generally greater than the number of concavities along its corresponding depth profile. A good example of this is shown in the lower right panel of Figure 3. The depicted object is the same as the one in the upper right panel, but the illumination is from the front rather than a peripheral direction. When a surface concavity faces directly toward the light source, the frequency of its luminance variations is doubled, producing two luminance concavities, and the deepest part of the surface concavity corresponds to a luminance maximum (see Langer & Bülthoff, 2000; Pentland, 1989; Todd, Egan, & Kallie, 2015). In this particular case, the prominent smooth occlusion contour makes it obvious that the central region of the surface is concave, but there are other situations where the number of apparent concavities is perceptually ambiguous for this type of stimulus.

The lower left panel of Figure 3 shows the same object as the upper left panel, but it is illuminated from two peripheral directions on the left and right. This creates a luminance valley that runs vertically from top to bottom, and it appears to us as a physical valley on the depicted surface. Note that the occlusion contour at the top and bottom indicates that that these regions are convex, but that seems to be overridden in this case by the pattern of luminance curvature. This observation suggests that spurious luminance concavities can create illusory perceptions of surface concavities.

Another important factor that can produce spurious luminance concavities is the pattern of surface curvature. Thus far, we have described the cross section through a surface as if it were a planar space curve with just one component of curvature. However, we can also think of a surface cross section as an elastic ribbon whose local orientations can vary in multiple directions. Figure 4 shows shaded images of ribbon surfaces with two types of curvature. The one on top has been subjected to a bending transformation such that all changes in the surface normal are confined to the plane of the central axis. This is the curvature that is depicted by the depth and luminance profiles in Figure 3. The surface on the bottom has been subjected to a twisting transformation such that all changes in the surface normal are perpendicular to the plane of the central axis. Note that both components of curvature can influence the pattern of shading along a surface cross section. In particular, the twist component can produce spurious luminance concavities in regions where the bending component has zero curvature or is uniformly convex.

The research described in the present article was designed to explore these phenomena in a more systematic manner. Its goals were threefold: first, to test the hypothesis that apparent surface concavities along an image cross section are primarily located within concave regions of the luminance profile; second, to determine the ability of human observers to compensate for phase shifts between corresponding depth and luminance profiles; and third, to investigate the extent to which spurious luminance concavities can produce illusory perceptions of surface curvature.

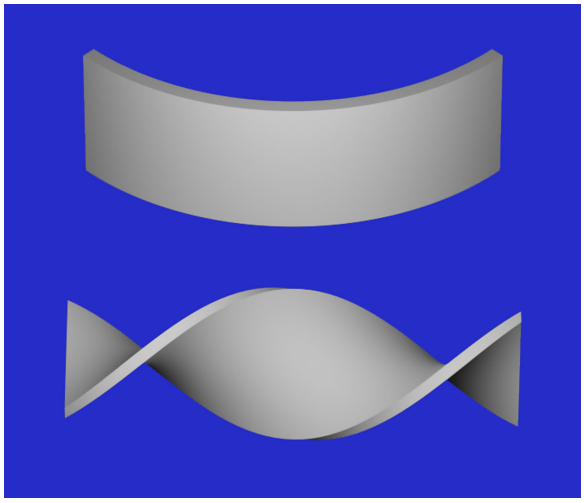

Figure 4. Ribbon surfaces with two types of curvature. The central axes of both surfaces are planar space curves. The top surface has undergone a bending transformation such that all changes in the surface normal are confined to the plane of the central axis. The bottom surface has undergone a twisting transformation such that all changes in the surface normal are perpendicular to the plane of the central axis. Note that both types of curvature can influence the shading along the central axis.

Experiment 1

In this experiment, observers were required to identify local concavities along designated horizontal cross sections of smoothly curved surfaces. All of the depicted materials were Lambertian, but they were presented with four different patterns of illumination.

Methods

Apparatus

The experiment was performed using a Dell Precision 3620 PC with an NVIDIA Quadro P4000 graphics card, running the Windows 10 operating system. The stimulus images were presented on a 10-bit, 27-in. LCD monitor with a spatial resolution of 2,560 × 1,440 pixels, a gamma of 2.2, and at a nominal luminance of 300 $\frac{cd}{m^2}$. The images had a spatial resolution of 1,200 × 1,200 pixels, and they were presented within a 22.4-cm × 22.4-cm region of the display screen. The displays

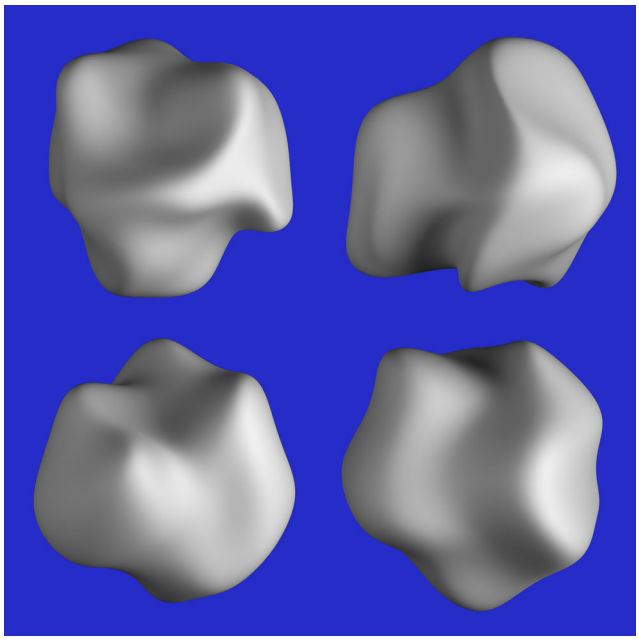

Figure 5. The four different objects used in the present experiments. All of them are illuminated primarily from the right.

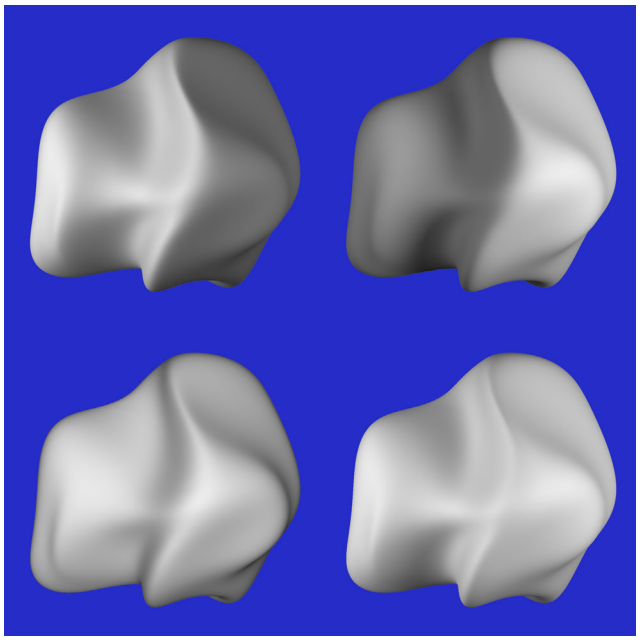

Figure 6. A single object with four different patterns of illumination used in the present experiment. Moving clockwise from the upper left, the object is illuminated primarily from the left, primarily from the right, equally from the left and right, and from the front.

were viewed monocularly at a distance of 70 cm, and head movements were restricted using a chinrest.

Stimuli

The stimuli were created using the Maxwell renderer, which is able to produce photo-realistic images with a space of possible materials that are all parameterized by physically measurable properties of real surfaces. Maxwell is an unbiased renderer in that it does not use heuristics to speed up rendering times at the cost of physical accuracy. All of the stimuli were rendered as 16-bit images at a high quality level, and their exposures were adjusted so that none of the intensity values was outside the range that is displayable on the 10-bit monitor. It is also important to emphasize that all of the renderings included the natural effects of cast shadows and indirect reflections, which have often been excluded in previous studies of the perception of 3D shape from shading. In order to compute luminance curvature for these stimuli, we needed to reduce the image noise as much as possible. This was facilitated by smoothing the images using Innobright's Altus denoising algorithm. These computations were performed using three networked workstations with a total of 96 Xeon cores.

All of the depicted objects had a Lambertian reflection function that scatters light equally in all directions. We used four different objects (see Figure 5) with four different horizontal cross sections and four different patterns of illumination in all possible combinations. Figure 6 shows one of the possible objects with all of the different patterns of illumination. Three of the illumination conditions used a standard photographic lighting setup with two area lights at a 55° angle relative to the line of sight on opposite sides of the depicted object. In the left illumination condition, the light on the left was five times more intense than the one on the right, and in the right illumination condition, the intensities of these lights were reversed. There was also an equal illumination condition in which both lights had the same intensity. Finally, a frontal illumination condition was included, in which a single small area light was positioned directly along the line of sight.

Procedure

There have been several previous studies in which observers have been asked to judge apparent depth changes along designated cross sections through an image. In one technique developed by Koenderink et al. (2001), a 3D surface is presented on a display screen and a line through the image is marked by a row of equally spaced dots (see also Egan et al., 2011; Liu & Todd, 2004; Thaler et al., 2007; Todd et al., 2004). Note that the projection of these dots onto the depicted surface defines a planar space curve. Adjacent to the image of a surface, an identical row of dots is presented against a blank background, and observers are asked to adjust the second row of dots so that it matches the apparent

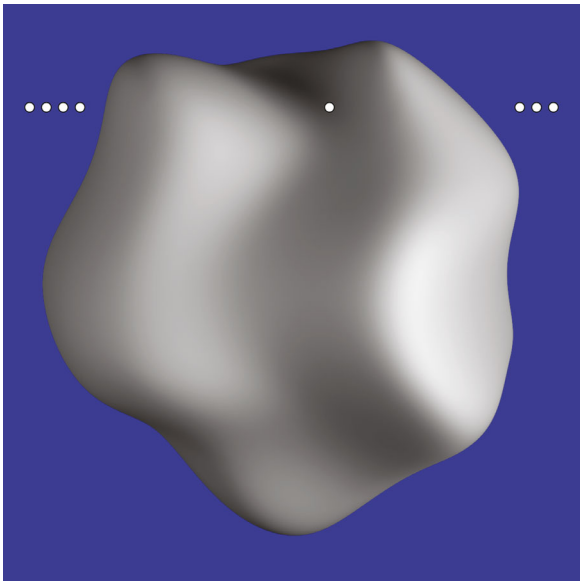

Figure 7. The task employed in Experiment 1. An image was presented with four small dots on each side to designate a particular cross section. Observers marked each apparent concavity along the cross section with one of the adjustable dots.

depth relief of the designated surface cross section. If these judgments are obtained for planar curves at multiple positions and orientations through the surface, it is possible to reconstruct a best-fitting smooth surface from those judgments (Koenderink et al., 2001).

In another related procedure, observers are presented with an image of a smooth surface, and a horizontal cross section through the surface is designated by a row of small dots on each side (see Figure 7). These dots can be individually moved along the designated cross section with a handheld mouse. Observers are required to identify any local depth extrema (minima or maxima) along the surface cross section and to mark each location with one of the adjustable dots (see Todd et al., 2004; Todd & Thaler, 2010). What is especially interesting about these tasks is that observers are able to isolate changes along a single planar space curve while ignoring the changes in any other direction. Thus, they can easily identify a local extremum in a designated direction even though it is not an extremum in an orthogonal direction.

A very similar procedure was employed in the present experiments. Observers were asked to identify surface concavities along designated surface cross sections and to mark the deepest part of each concavity with one of the adjustable dots. Note that the definition of "deepest" in this context is defined by the orientation of the concavity rather than its distance from the point of observation (see Figure 8). Before participating in the actual experiment, the observers were trained using a pencil-and-paper task. We first explained the concept

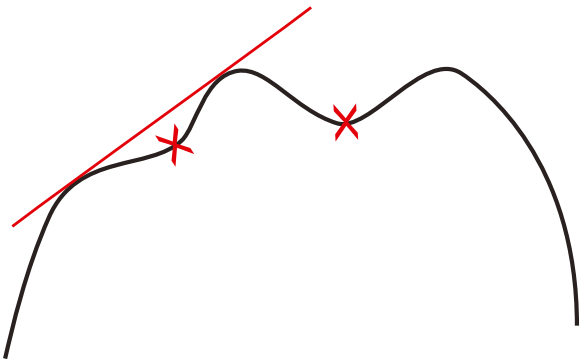

Figure 8. Observers were instructed to mark the deepest point within a concavity relative to a line that is tangent to both of its boundaries. They practiced this task on simple line drawings like the one shown here by marking an X at the appropriate points. Note that the concavity on the right coincides with a depth extremum, but that is not the case for the one on the left.

of a one-dimensional (1D) curvature and then showed them a series of smooth curves like the one in Figure 8. They were required to mark the deepest part of each concavity as defined by the instructions by drawing an X on it. Note in Figure 8 that the depicted curve has two concavities. The deepest part of the one on the right is located at a depth extremum, but that is not the case for the one on the left. All of the observers demonstrated that they understood the instructions by marking the correct points on this preliminary task. Observers were specifically instructed to judge the 1D curvature along the designated cross sections and to ignore any components of curvature in other directions.

Ten observers judged each of the 64 possible scan lines five times each over four experimental sessions. During any given experimental session, each object was presented with just a single pattern of illumination, so that judgments of a given object with one illumination could not influence judgments of the same object with a different illumination presented shortly thereafter. All of the observers had normal (or corrected-to-normal) visual acuity.

Results

The first step for analyzing the data was to compute the surface and luminance curvatures along each surface cross section that observers were asked to judge. This was achieved by fitting a second-order patch centered on each pixel in an image using a least squares procedure. The patch was defined by the following equation: $Z = aX^2 + bY^2 + cXY + dX + eY$, where X and Y are the horizontal and vertical coordinates of a given pixel and Z is the intensity of the luminance surface at each pixel location. From the coefficients of the best-fitting patch, we could analytically compute

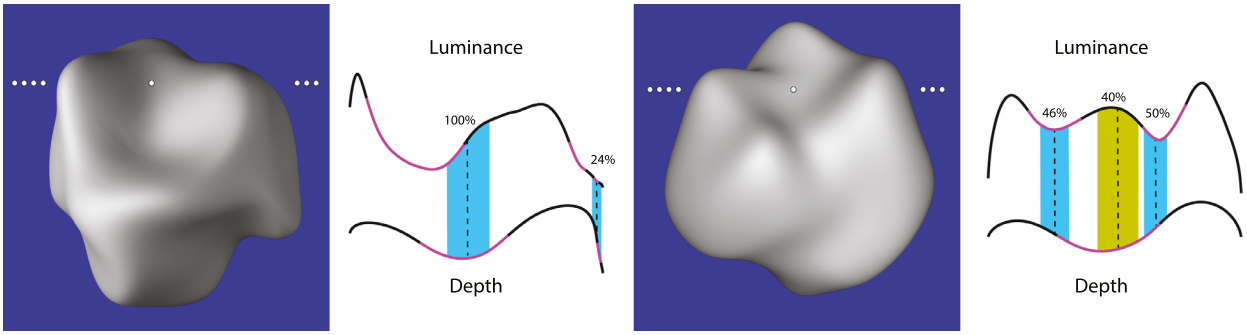

Figure 9. Some patterns of response for Experiment 1. Each panel contains a shaded image on the left with a designated surface cross section that observers were asked to judge and a typical pattern of response for that cross section. The right side of each panel shows the depth and luminance profiles along the designated surface cross section. Convex regions along each curve are colored black, and concave regions are colored magenta. The dashed black lines show the mean locations where observers' responses were clustered, the cyan or yellow band around each line shows ± two standard deviations of the response distribution, and the number just above each line shows the percentage of trials where that region was marked. Note in the left panel that there is a phase shift between the depth and luminance profiles and that observers are able to compensate for that. The image on the right is perceptually ambiguous. Some observers see a single concavity centered on the luminance maximum, whereas as others see two distinct concavities centered on the luminance minima.

all aspects of curvature, including mean curvature, Gaussian curvature, and the principal directions of curvature. For these particular stimuli, a patch that spanned 41 × 41 pixels yielded excellent results such that the r^2 values were almost always in excess of 0.95. For regions near the occlusion boundary of an object, all background pixels were excluded from the least squares analysis. The curvatures of the depicted surfaces (i.e., the ground truth) were computed in exactly the same way. In that case, the computations were performed on a high-resolution 16-bit depth map of the surface that was smoothed with a denoising algorithm. Curvatures were computed by fitting a second-order patch centered on each pixel to determine the best-fitting parameters of that patch.

Let us now consider some specific examples of the observers' responses. Figure 9 is divided into two panels, each of which contains a shaded image on the left and a typical pattern of response along a single horizontal cross section that observers were asked to judge. The right side of each panel shows the depth and luminance profiles along that cross section. Convex regions along each curve are colored black, and concave regions are colored magenta. The dashed black lines connecting those curves show the mean locations of observers' responses as determined by an optimal *k*-means cluster analysis; the bands around each line show ± two standard deviations of the response distribution; the number just above each line shows the percentage of possible responses where that region was marked.

The left panel of Figure 9 shows the generic case where the depicted object is illuminated from a peripheral direction. Note that the left side of the object is much brighter than the right side, which indicates that it is primarily illuminated from the left. The luminance profile has the same spatial frequency as the depth profile, but it is phase shifted to the left. When observers were asked to mark the deepest point of the surface concavity, they were able to take that phase shift into account. That is to say, their responses were clustered near an inflection point on the luminance profile along a sidewall that faces toward the direction of illumination, and that region was marked on 100% of the possible trials. Many of these responses were within the concave region of the luminance profile, although others were located just across the boundary within a convex region. Note that there are two other small luminance concavities on the right side of the cross section in this figure. One of these corresponds to an extremely shallow concavity that is barely detectable in the surface depth profile, but that region was still marked on 24% of the possible trials.

The right panel of Figure 9 shows a different object cross section with frontal illumination that produces a more complex relationship between its depth and luminance profiles. As was pointed out by Pentland (1989), frontal illumination can cause the spatial frequency of the luminance profile to be double that of the depth profile, and the deepest part of a surface concavity in that case often corresponds to a local luminance maximum (Langer & Bülthoff, 2000; Todd et al., 2015). Note that the depth profile has two peaks and one trough, whereas the luminance profile has three peaks and two troughs. Our results show that this particular stimulus is perceptually ambiguous. Some observers report that there is a single surface concavity along the designated cross section (which is

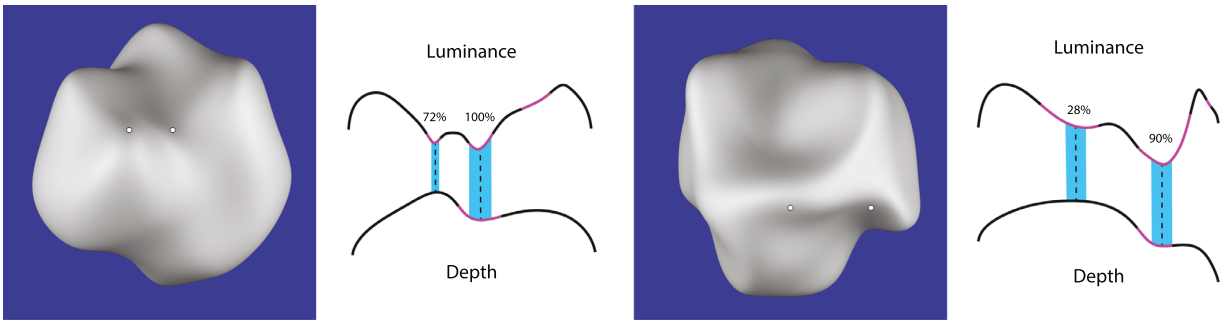

Figure 10. Examples of spurious luminance concavities from Experiment 1 that do not correspond to a surface concavity. Both of these images depict surfaces that are illuminated equally from the left and right. This causes spurious luminance concavities to occur in regions where the surface depth is a local minimum, and these regions are often mistakenly judged as surface concavities along the designated surface cross section.

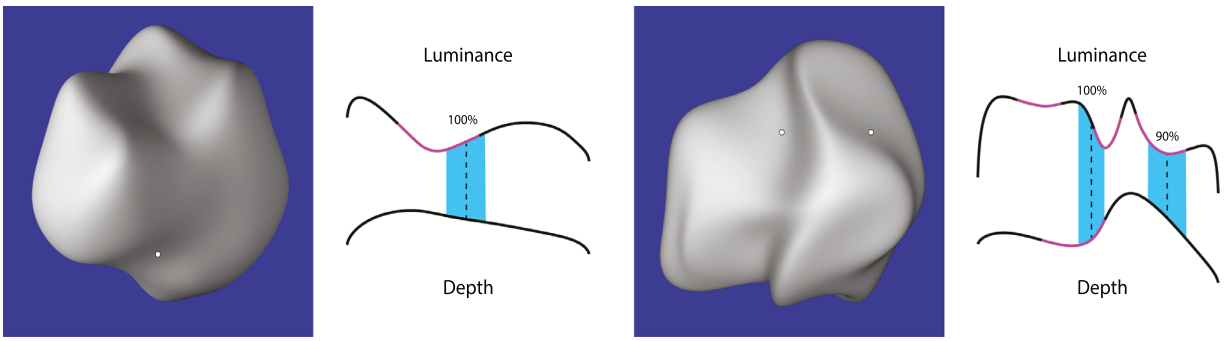

Figure 11. Examples of spurious luminance concavities from Experiment 1 that do not correspond to a surface concavity. The designated cross sections in these images both contain a twist in the pattern of curvature that causes spurious luminance concavities in regions where the vertical component of the surface normal is a local extremum in the horizontal direction. These regions are often mistakenly judged as surface concavities along the designated surface cross section.

the physically correct response), and their settings are typically clustered near a local luminance maximum. This is represented in the figure by a yellow band. Other observers report that there are two distinct concavities along the cross section that appear to be separated by a convex region, and those responses are typically clustered near the luminance minima. These are represented in the figure as cyan bands.

It is important to point out that the presence of a luminance concavity does not guarantee that a corresponding region of the depth profile will contain a concavity. It is possible to have spurious luminance concavities along a surface cross section in which the depth profile is uniformly convex. If observers use luminance concavities as a source of information about surface concavities, they are likely to produce false alarm responses in those regions. Figure 10 shows two examples of spurious luminance concavities from the equal illumination condition, which produced two to three times more spurious concavities than any of the other conditions. In both examples, there is a spurious luminance concavity on the left that corresponds to a peak of the surface depth profile. Observers did indeed mark these regions as apparent surface concavities on 72% and 28% of the trials, respectively. It is also interesting to note in this figure that there is no phase shift in the observers' responses, which are clustered around the lowest point on all of the luminance concavities. This is likely because the two ends of the designated cross section both have the same luminance, thus indicating that there is no peripheral skew to the direction of illumination.

Two other examples of spurious luminance concavities are shown in Figure 11. These are both caused by twist in the pattern of curvature along the designated surface scan line. That is to say, the vertical component of the surface normal changes systematically with its horizontal position along the designated cross section (see Figure 12). This causes local reductions of luminance, which are perceptually interpreted as a surface concavities. In the left panel of Figure 11, the depicted cross section was uniformly

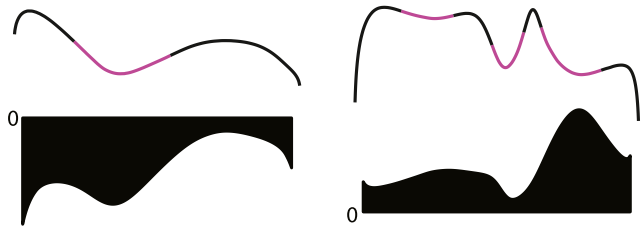

Figure 12. The curves on the top show the luminance profiles from Figure 11. The ones on the bottom show the vertical component of the surface normal along those cross sections relative to zero (i.e., the twist). Note how the spurious luminance concavities are aligned with local extrema in the vertical slant of the surface.

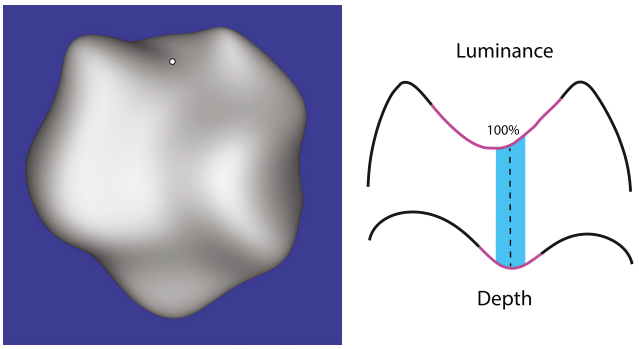

Figure 13. An image of a surface with frontal illumination. Frequency doubling does not occur along the designated cross section because of twist in the surface curvature. As the depth of the surface increases near the center of the scan line, there is also an increase in the vertical tilt of the surface that causes the shading to become darker.

convex, but the twist of the surface normals produced a spurious luminance concavity that was marked as a physical concavity on 100% of the trials. In the right panel, the spurious luminance concavity was marked on 90% of the trials.

Another interesting effect of twist is shown in Figure 13. The left panel shows a surface with frontal illumination like the right panel of Figure 9. Horizontal cross sections through the center of the object exhibit the frequency doubling that is characteristic of frontal illumination, but the cross section that is marked near the top does not. As the depth of the surface increases near the center of the scan line, there is also an increase in the vertical tilt of the surface that causes the shading to become darker. This is the result of twist in the pattern of surface curvature along the designated surface cross section.

It is interesting to note that observers marked actual surface concavities on 74% of the trials, but spurious

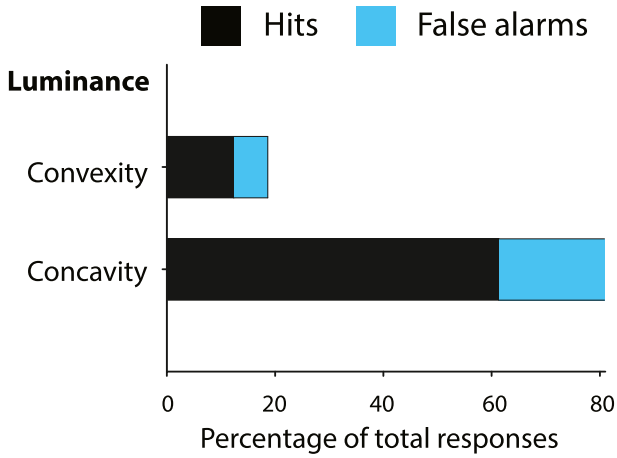

Figure 14. The overall results for Lambertian surfaces in Experiment 1.

luminance concavities were marked on only 44% of the trials. One possible reason for this is that the spurious luminance concavities tended to have lower amplitudes than those that are associated with concavities in the depicted surface. There were large individual differences in the number of apparent concavities that were marked. Some observers consistently marked tiny concavities that are barely visible, whereas others were more conservative. There may also have been other sources of information that competed with luminance curvature to identify which luminance concavities were likely to be spurious.

Let us now consider the overall pattern of observers' responses. An analysis of the different possible image cross sections revealed that 57% of the pixels were in convex regions of the luminance profiles and 43% were in concave regions. If observers had responded randomly, then that same relative proportion of concavities and convexities should also be reflected in the pattern of responses, but that is not what occurred. Out of 4,372 total responses, only 19% were in convex regions of the luminance profiles and 81% were in concave regions. The stacked bar chart in Figure 14 provides a breakdown of the hits and false alarms. Note that most observers' responses were in concave regions of the luminance profile. Some of those responses (about 20%) were false alarms in spurious luminance concavities that did not correspond to a surface concavity. Of the responses in convex regions of the luminance profile, roughly two thirds were hits that fell within regions where the lighting was aligned with the central axis of a concavity so that its deepest part is a luminance maximum (Langer & Bülthoff, 2000; Pentland, 1989; Todd, Egan, & Kallie, 2015) or a convex region that is just across the boundary of a concave region that faces toward the direction of illumination.

Experiment 2

Most of the early research on the perception of shape from shading was restricted to surfaces with Lambertian reflectance functions, but that has gradually changed with the development of more sophisticated computer rendering models (e.g., Fleming et al., 2004; Khang et al., 2007; Marlow & Anderson, 2021; Mooney & Anderson, 2014; Nefs et al., 2006; Pont & te Pas, 2006; Sawayama & Nishida, 2018; Todd et al., 2014; Wijntjes et al., 2012). Experiment 2 was designed to determine if human observers can reliably identify surface concavities from the patterns of reflection on non-Lambertian materials and whether those judgments are consistent with the results obtained with Lambertian materials.

Methods

The apparatus, procedure, and stimulus objects were the same as in Experiment 1, and the stimuli were judged by the same 10 observers. The most important differences were that the simulated objects were all illuminated primarily from the left, and they all had non-Lambertian reflectance functions like the ones depicted in Figure 15, including glossy paint, black velvet, satin cloth, and wax. The glossy paint material had an additive combination of diffuse and specular components. The black velvet material was purely specular with grazing angle highlights. The satin cloth material was designed to produce anisotropic reflections. These occur on surfaces with microgrooves such as brushed metal or woven cloth that run in one dominant direction. This causes light to reflect in a specular way in the direction of the grooves and in a more diffuse way in the direction perpendicular to the grooves. Finally, we also included a wax material for which light is scattered inside the volume of an object.

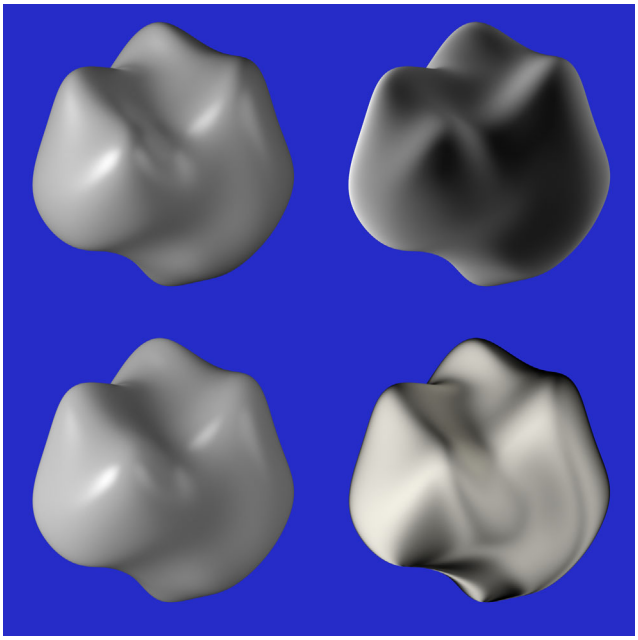

Figure 15. A single object with four different reflectance functions used in Experiment 2. Moving clockwise from the upper left, the depicted materials are glossy paint, velvet cloth, satin cloth, and wax.

Results

Figure 16 shows two examples of the same object with different materials. The one on the left depicts a velvet material with grazing angle highlights. The luminance profile along the designated cross section has two prominent concavities. One of these corresponds to an actual surface concavity, and that one was marked on 100% of the possible trials, although a few of those responses were just outside the concave region of the depth profile. There is also a spurious luminance concavity on the right that was marked on 80% of the trials. These responses were all categorized as false alarms.

The right panel of Figure 16 shows the same object as on the left with a glossy paint material. The luminance profile in that case has five distinct concavities. One of these is much larger than the others and is phase shifted to the left of an actual surface concavity. That region was marked on 100% of the trials, and all of those were hits that fell within the concave region of the depth profile. The are also four spurious concavities that are much smaller, but only one had a response rate over 50%. It is important to keep in mind that these displays were illuminated by a bright light on the left and a dimmer one on the right. The luminance concavity that is second from the right is flanked by specular highlights from both light sources. This creates the appearance of a small vertical groove in the surface that observers marked on 54% of the trials.

The stacked bar chart in Figure 17 provides a summary of the overall pattern of performance. Note that the results are remarkably similar to those from Experiment 1, even though the surface reflectance functions were quite different. Most observers' responses were in concave regions of the luminance profile, and almost all of the remaining responses were either just outside the boundary of a concave region or at a local luminance maximum where the lighting direction was aligned with the central axis of a surface concavity. These findings are quite surprising because they provide strong evidence that the perceived qualitative shapes of Lambertian and non-Lambertian surfaces may be largely based on the same optical information.

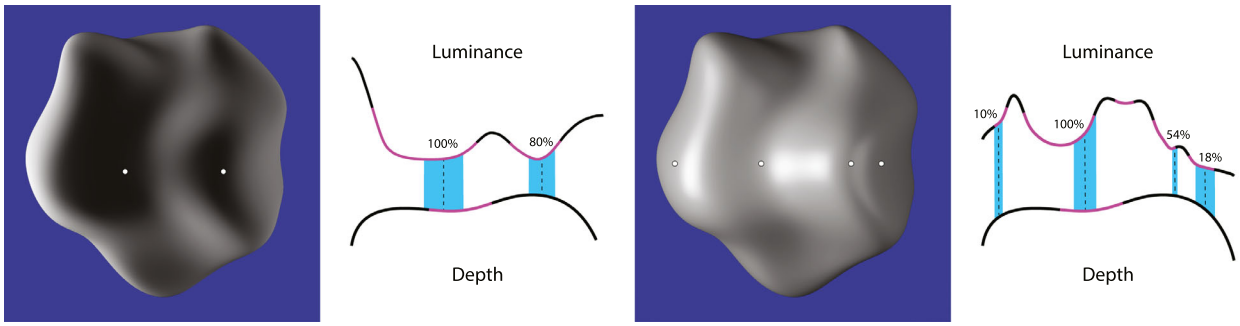

Figure 16. Patterns of response for a single object with two different reflectance functions. The left image depicts black velvet cloth, and the right one depicts glossy paint.

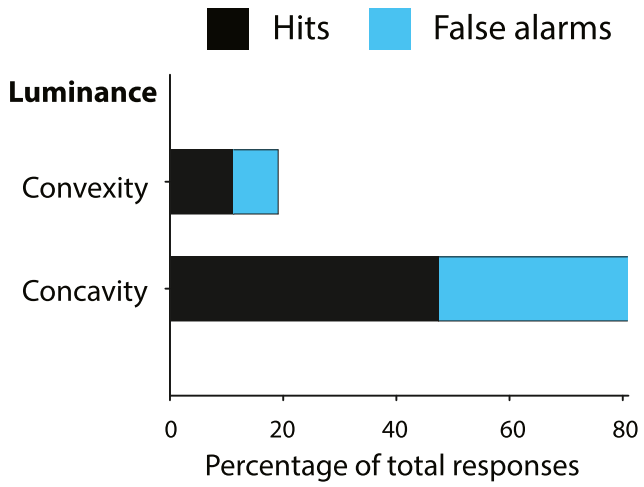

Figure 17. The overall results for non-Lambertian surfaces in Experiment 2.

Note in Figures 14 and 17 that the false alarm rate increased significantly in Experiment 2 relative to Experiment 1. This is likely because the non-Lambertian materials produced a greater number of spurious luminance concavities. Figure 11 shows several examples for velvet and glossy paint materials. As in Experiment 1, the spurious luminance concavities were marked on a smaller percentage of trials (47%) than those that corresponded to an actual surface concavity (74%). This again suggests that those judgments were influenced by some other source of information in addition to luminance curvature that may have helped to identify which luminance concavities were likely to be spurious.

Discussion

An obvious difficulty for the perception of 3D shape from shading is that the pattern of luminance that reflects off an object toward the point of observation can change dramatically over variations in the pattern of illumination or the surface reflectance function. There is abundant research to demonstrate that observers' perceptions of 3D shape can be distorted by these changes (see Anderson & Marlow, 2022, for a review), but the precise nature of those distortions can be difficult to pin down. One experiment by Egan and Todd (2015) tried to do that using a local orientation adjustment task on stimuli like the ones in the present study with varying patterns of illumination. They analyzed the data using an algorithm developed by Koenderink, van Doorn, and Kappers (1992, 1995) to calculate the best-fitting surface that is maximally consistent with the overall pattern of an observer's judgments in each condition. These reconstructed surfaces were then correlated with the ground truth. The results revealed that about 60% of the error was due to a simple scaling in depth, in which observers consistently underestimated the overall relief of the objects. Another 28% of the error was due to shear toward the direction of the light source. They also did a test–retest correlation across different blocks, and the residual variance in that case was just 2%. That leaves another 10% of systematic error that cannot be explained by an affine transformation between observers' perceptions and the ground truth (see also Nefs et al., 2005). It is this qualitative nonaffine structure that was the focus of the present investigation.

An important assumption of our research is that the appearance of shape constancy must be based on the qualitative properties of objects that are invariant over affine transformations. These properties are more stable over changing viewing conditions than traditional methods of representing shapes such as local surface depths or orientations, and the visual information by which they are specified requires an analysis of the second-order differential structure of the luminance field. This approach is exemplified in one of the earliest analyses of shape form shading by Koenderink and van Doorn (1980). They studied the properties of saddle points in the luminance field that are not local maxima or minima but where the first spatial derivative

is zero in all directions. Saddle points can also be defined as regions where isointensity contours cross one another. For Lambertian surfaces with homogeneous illumination, saddle points in the luminance field will always correspond to parabolic points on a surface where one of the principal curvatures is zero. A similar analysis has been developed more recently by Kunsberg and Zucker (2018, 2021). They exploit saddle points in the luminance field and local maxima of luminance gradients to construct critical contours in an image that define the boundaries of bumps and dimples.

Our examination of qualitative structure looked at concavities and convexities of the luminance field along individual surface cross sections. Observers were asked to mark the deepest point on each perceived concavity along a cross section with an adjustable dot. One possible strategy for performing this task is to adopt the "darker-is-deeper" heuristic proposed by Langer and Zucker (1994). According to that hypothesis, observers should always pick points that are local luminance minima. This did occur on some trials, especially in the equal illumination condition. However, there were many other trials where observers marked local luminance maxima, and most of those responses were consistent the ground truth. The deepest part of a concavity will be a local luminance maximum whenever the concavity faces directly toward the primary source of illumination (see Langer & Bülthoff, 2000; Todd, Egan, & Kallie, 2015). When that occurred in the present experiments, the results indicated that these regions can be perceptually multistable. Observers sometimes mark a local luminance maximum, which is the correct response, but they can also mark the two local luminance minima on each side, which adds an additional concavity relative to the ground truth. That is an example of a nonaffine perceptual distortion.

For peripheral directions of illumination, the luminance profile is phase shifted relative to the depth profile, and the deepest part of a concavity will be somewhere in between a local luminance minimum and a local luminance maximum. The results indicate that observers are able to compensate for these phase shifts to help identify the correct locations of surface concavities (see also Todd & Reichel, 1989). A likely source of information to make that possible is the relative luminance along smooth occlusion contours. For example, if the right side of a surface cross section is much brighter than the left, then the direction of illumination is skewed to the right, and the deepest part of a concavity is located to the left of the luminance minimum. An exception to this occurs in the equal illumination condition. Because that condition balances the luminance on both sides of a surface cross section, there is no reason to compensate for peripheral illumination, and observers' responses tend to follow the darker-is-deeper rule (e.g., see Figure 10).

Our primary interest in conducting these experiments was to discover how observers' responses would be affected by the presence of spurious luminance concavities that do not correspond to a surface concavity. These can occur when an object is illuminated from opposite sides at the same time or when there is twist in the curvature of the surface along a designated cross section. If luminance concavities are an important source of information about the presence of surface concavities, then spurious luminance concavities should produce false alarm responses in our concavity detection task. This prediction was clearly confirmed by the results. With Lambertian surfaces in Experiment 1, 20% of the responses were false alarms to spurious concavities, and that rate increased to 32% for non-Lambertian surfaces in Experiment 2. This provides compelling empirical evidence that luminance concavities are a perceptually useful source of information about the presence of surface concavities.

There are other aspects of our data to suggest that the effects of luminance curvature can be modulated by other sources of information. For example, the spurious luminance concavities were marked at a much lower rate than those that corresponded to an actual surface concavity. Some of that may be due to the relative widths and magnitudes of those concavities, but we suspect there are other factors involved. One likely source of information for modulating luminance curvature is an object's smooth occlusion contour (Egan & Todd, 2015; Marlow, Mooney, & Anderson, 2019; Mooney, Marlow, & Anderson, 2019). This information is especially salient for regions that are close to the contour, but its impact is likely to diminish for interior regions that are farther away.

One possible criticism of this research is that the judgments were limited to 1D cross sections of a surface and cannot therefore provide useful information about the overall shape of an object. However, this type of criticism ignores the wealth of information that can be obtained from many different measurements in neighboring regions (e.g., see Koenderink et al., 2001) For example, Figure 18 shows a series of concavity judgments along numerous horizontal cross sections of an object. The cyan curve that connects the judged points reveals a systematic pattern of apparent negative curvature that forms an s-shaped valley along the entire vertical extent of the surface. Hoffman and Richards (1984) have argued that contours connecting local negative curvature extrema are ideal candidates for segmenting smooth surfaces into parts. Hertzmann (2020) has proposed a related hypothesis that the contours used to depict smooth surfaces in line drawings often correspond to valleys of negative curvature in the luminance field. It will be interesting to see in future research if that finding holds up over variations in surface materials or the pattern of illumination.

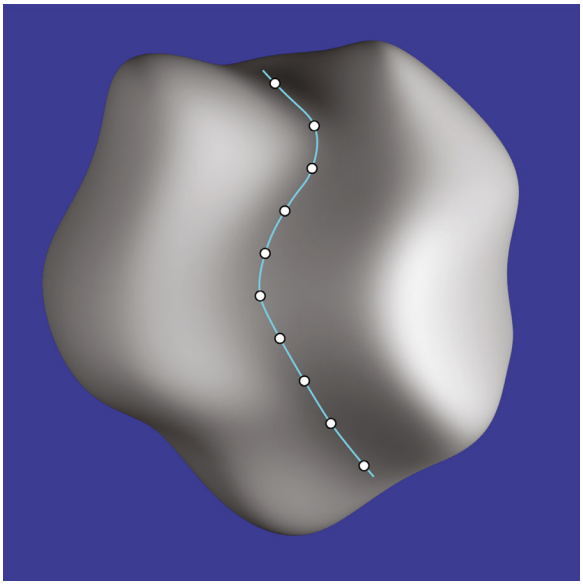

Figure 18. Concavity judgments from a series of horizontal cross sections reveal an s-shaped valley along the entire vertical extent of the depicted object.

An implicit assumption of our computational analysis is that the convex regions of a luminance profile have negative second spatial derivatives and that concave regions have positive second spatial derivatives. It follows from this assumption that the perceived sign of relief should be reversed if the intensity gradients within an image are inverted. Figure 19 shows a pair of images to test that prediction. The one on the left is a stimulus from Experiment 1, and the one on the right is a negative version of that. Two corresponding points in each image are marked by small dots. Note in the left panel that the two dots appear to be in concave regions of the surface and that the second derivative of the luminance field in a horizontal direction is positive. The dots in the right panel also appear to be in concave surface regions, but the second derivatives of shading in those regions are negative. Indeed, the reversal of shading in this case seems to have a negligible influence on the overall perceived shape of the object.

One possible explanation of this involves a simple cue conflict. The smooth occlusion contours adjacent to the two dots provide information that neighboring surface regions are concave, and that overrides the information provided by the luminance curvature. However, that cannot be the whole story. Smooth occlusion contours are only informative about surface regions in their immediate local neighborhoods, but negative images cause shading gradients to be reversed over the entire surface. This suggests that some other source of information may trigger a reinterpretation of luminance curvature so that positive second derivatives are associated with surface convexities, and negative second derivatives are associated with surface concavities.

A key to understanding how this might be possible is that positive and negative images are generally easy to distinguish from one another. When viewing a real surface (or a positive image of one), the surface curvature perpendicular to a smooth occlusion contour is always convex in its immediate local neighborhood on the attached side (Koenderink, 1984; Koenderink & van Doorn, 1982). The luminance curvature perpendicular to a smooth occlusion contour is most often convex as well, but that is not guaranteed. To get a better sense of the overall statistics, we examined the luminance curvature at the ends of the 128 cross sections observers were asked to judge, which provided a sample of 256 regions in the immediate vicinity of a smooth occlusion contour. If we exclude the black velvet conditions, over 97% of the adjacent regions had a convex luminance curvature. Thus, if the distribution of curvatures around a smooth occlusion contour is

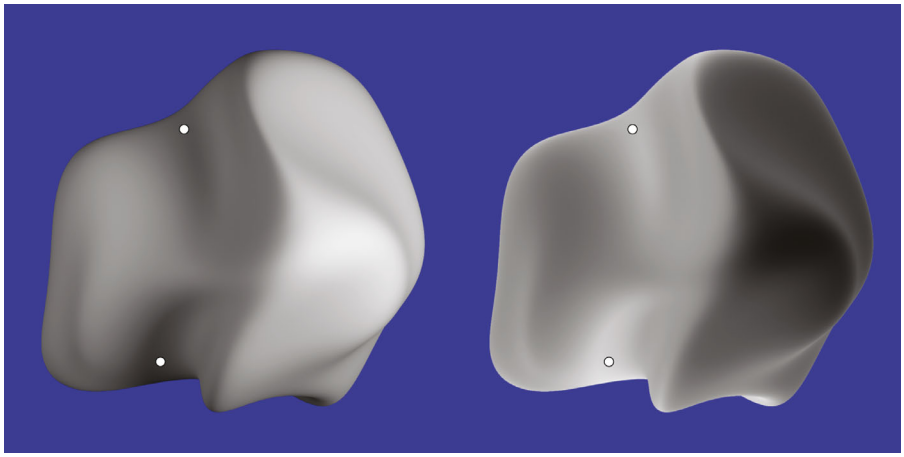

Figure 19. A positive image of a smooth surface (left) and a negative image (right) produced by inverting the shading gradients from the one on the left. Corresponding regions in these images are marked by small dots.

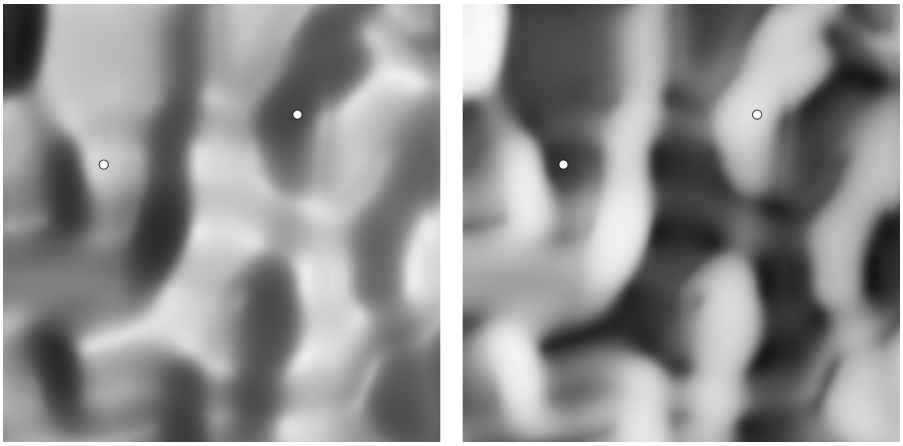

Figure 20. Positive and negative images of a random noise surface. Corresponding regions in these images are marked by small dots. Note that the apparent relief for the negative image on the right is reversed relative to the positive image on the left. The region marked by the leftmost dot appears convex in the positive image and concave in the negative image. Conversely, the region marked by the rightmost dot appears concave in the positive image and convex in the negative image.

overwhelmingly convex, that provides strong evidence that the luminance pattern is a positive image of a 3D surface and that negative second derivatives of luminance are associated with surface convexities. If the distribution is overwhelmingly concave, then the luminance pattern is likely to be a negative image of a 3D surface, and negative second derivatives of luminance are associated with surface concavities. The black velvet conditions were outliers in this analysis because 78% of the sampled regions had a concave luminance curvature. However, those conditions could easily be mistaken for negative images (e.g., examine the black velvet surfaces in Figures 15 and 16).

Figure 20 is intended to show what happens when surfaces are viewed without any smooth occlusion contours at all. The left panel depicts a positive image of a random noise surface illuminated from the right, and two of its regions are marked by small dots. A negative image of the same surface is shown in the right panel with the same pair of corresponding dots. Note that the apparent sign of surface relief is completely reversed in that case. The region marked by the leftmost dot appears convex in the positive image and concave in the negative one. Conversely, the region marked by the rightmost dot appears concave in the positive image and convex in the negative one.

In the absence of smooth occlusion contours, it is difficult to discern the direction of illumination or to distinguish between positive and negative images. Because much of the information about 3D shape has been removed, observers can only rely on ad hoc heuristics to distinguish concavities from convexities. This could involve a bias to perceive surfaces such that apparent depth increases with height in the visual field (Langer & Bülthoff, 2001; Reichel & Todd, 1990) or a bias to perceive surfaces as convex rather than concave (Langer & Bülthoff, 2001; Liu & Todd, 2004). For the example in Figure 14, observers seem to adopt the "darker-is-deeper" heuristic of Langer and Zucker (1994), even though that is likely to produce inaccurate responses (see Langer & Bülthoff , 2000; Todd, Egan, & Kallie, 2015). The key thing to note in all of these examples is that the overall pattern of concavities and convexities remains invariant. All that is affected is whether surface concavities are defined visually by positive or negative second derivatives of the luminance field.

Conclusions

When human observers attempt to verbally describe 3D surfaces, they typically identify salient topographic features, such as bumps, dimples, ridges, valleys, and saddles, and how those features are arranged with respect to one another. These qualitative descriptions are defined by the pattern of surface concavities and convexities, and the visual information on which they are based involves the second-order differential structure of the luminance field (Koenderink & van Doorn, 1980; Kunsberg & Zucker, 2018, 2021). The research described in the present experiments investigated perceived qualitative shape by asking observers to identify concavities along surface cross sections. The results revealed that most responses were in concave regions of the luminance profiles, although they were often shifted in phase relative to the curvature of the depicted surfaces. This pattern of performance was surprisingly robust over large changes in the pattern of illumination or surface material properties. Our analysis predicts that observers should make false

alarm responses in regions where a luminance concavity does not correspond to a surface concavity, and our empirical results confirmed that prediction. We also showed how the interaction of luminance curvature with smooth occlusion contours can account for the perception of 3D shape in negative images for which all luminance gradients are reversed.

Keywords: shape from shading, luminance curvature, shape constancy

Acknowledgments

Supported by a grants from the National Science Foundation (BCS-1849418, BCS-2238179, BCS-2238180).

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

References

Anderson, B. L., & Marlow, P. J. (2022). Perceiving the shape and material properties of 3D surfaces. *Trends in Cognitive Science, 27*, 98–110.

Belhumeur, P. N., Kriegman, D. J., & Yuille, A. L. (1999). The bas-relief ambiguity. *International Journal of Computer Vision, 35*(1), 33–44.

Caniard, F., & Fleming, R. (2007). Distortion in 3D shape estimation with changes in illumination. *Proceedings of the 4th Symposium on Applied Perception in Graphics and Visualization, 253*, 99–105.

Christou, C. G., & Koenderink, J. J. (1997). Light source dependence in shape from shading. *Vision Research, 37*(11), 1441–1449.

Curran, W., & Johnston, I. A. (1996). The effect of illuminant position on perceived curvature. *Vision Research, 36*(10), 1399–1410.

Egan, E. J. L., Todd, J. T., & Phillips, F. (2011). The perception of 3D shape from planar cut contours. *Journal of Vision, 11*(12):15, 1–13, doi: 10.1167/11.12.15.

Egan, E. J. L., & Todd, J. T. (2015). The effects of smooth occlusions and directions of illumination on the visual perception of 3-D shape from shading. *Journal of Vision, 15*(2):24, 1–11, doi: 10.1167/15.2.24.

Fleming, R. W., Torralba, A., & Adelson, E. H. (2004). Specular reflections and the perception of shape. *Journal of Vision, 4*, 798–820, doi: 10.1167/4.9.10.

Hoffman, D. D., & Richards, W. A. (1984). Parts of recognition. *Cognition, 18*(1–3), 65–96, doi: 10.1016/0010-0277(84)90022-2.

Horn, B. (1975). Obtaining shape from shading information. In P. Winston (Ed.), *The psychology of computer vision* (pp. 115–155). New York, NY: McGraw-Hill.

Hertzmann, A. (2020). Why do line drawings work? A realism hypothesis. *Perception, 49*(4), 439–451.

Ikeuchi, K., & Horn, B. K. P. (1981). Numerical shape from shading and occluding boundaries. *Artificial Intelligence, 17*(1–3), 141–184.

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

Koenderink, J. J., & van Doorn, A. J. (1980). Photometric invariants related to solid shape. *Optica Acta: international Journal of Optics, 27*(7), 981–996.

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

Koenderink, J. J., van Doorn, A. J., Christou, C., & Lappin, J. S. (1996a). Shape constancy in pictorial relief. *Perception, 25*, 155–164.

Koenderink, J. J., van Doorn, A. J., Christou, C., & Lappin, J. S. (1996b). Perturbation study of shading in pictures. *Perception, 25*, 1009–1026.

Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. L. (1992). Surface perception in pictures. *Perception & Psychophysics, 52*, 487–496.

Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. L. (1995). Depth relief. *Perception, 24*(1), 115–126.

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

Khang, B. G., Koenderink, J. J., & Kappers, A. M. L. (2007). Shape from shading from images rendered with various surface types and light fields. *Perception, 36*(8), 1191–1213, doi: 10.1068/p5807.

Kunsberg, B., & Zucker, S. W. (2018). Critical contours: an invariant linking image flow with salient surface organization. *SIAM Journal on Imaging Sciences, 11*(3), 1849–1877.

Kunsberg, B., & Zucker, S. W. (2021). From boundaries to bumps: when closed (extremal) contours are critical. *Journal of Vision, 21*(13):7, 1–27, doi: 10.1167/jov.21.13.7.

Langer, M. S., & Bülthoff, H. H. (2000). Depth discrimination from shading under diffuse lighting. *Perception, 29*(6), 649–660, doi: 10.1068/p3060.

Langer, M. S., & Bülthoff, H. H. (2001). A prior for local convexity in local shape from shading. *Perception, 30*, 403–410.

Langer, M. S., & Zucker, S. (1994). Shape from shading on a cloudy day. *Journal of the Optical Society of America A, 11*(2), 467–478.

Liu, B., & Todd, J. T. (2004). Perceptual biases in the interpretation of 3D shape from shading. *Vision Research, 44*(18), 2135–2145, doi: 10.1016/j.visres.2004.03.024.

Lee, C. H., & Rosenfeld, A. (1985). Improved methods of estimating shape from shading using light source coordinate system. *Artificial Intelligence, 26*, 125–143.

Marlow, P., & Anderson, B. (2021). The cospecification of the shape and material properties of light permeable materials. *Proceedings of the National Academy of Sciences of the United States of America, 118*(14), e2024798118.

Marlow, P., Mooney, S., & Anderson, B. (2019). Photogeometric cues to perceived surface shading. *Current Biology, 29*(2), 306–311.

Mooney, S. W. J., & Anderson, B. L. (2014). Specular image structure modulates the perception of three dimensional shape. *Current Biology, 24*, 2737–2742.

Mooney, S., Marlow, P., & Anderson, B. (2019). The perception and misperception of optical defocus, shading, and shape. *eLife, 8*, 1–23.

Nefs, H. T., Koenderink, J. J., & Kappers, A. M. L. (2005). The influence of illumination direction on the pictorial reliefs of Lambertian surfaces. *Perception, 34*(3), 275–87, doi: 10.1068/p5179.

Nefs, H. T., Koenderink, J. J., & Kappers, A. M. L. (2006). Shape-from-shading for matte and glossy objects. *Acta Psychologica, 121*(3), 297–316.

Pentland, A. P. (1984). Local shading analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 6*(2), 170–187, doi: 10.1109/tpami.1984.4767501.

Pentland, A. P. (1989). Shape information from shading: a theory about human perception. *Spatial Vision, 4*(2–3), 165–182.

Pont, S. C., & te Pas, S. F. (2006). Material illumination ambiguities and the perception of solid objects. *Perception, 35*, 1331–1350.

Reichel, F. D., & Todd, J. T. (1990). Perceived depth inversion of smoothly curved surfaces due to image orientation. *Journal of Experimental Psychology: human Perception and Performance, 16*, 653–664.

Sawayama, M., & Nishida, S. (2018). Material and shape perception based on two types of intensity gradient information. *PLoS Computational Biology, 14*(4), e1006061, doi: 10.1371/journal.pcbi.1006061.

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

Todd, J. T., Egan, E. J. L., & Kallie, C. S. (2015). The darker-is-deeper heuristic for the perception of 3D shape from shading: is it perceptually or ecologically valid? *Journal of Vision, 15*(11):24, 1–9, doi: 10.1167/15.15.2.

Todd, J. T., Egan, J. L., & Phillips, F. (2014). Is the perception of 3D shape from shading based on assumed reflectance and illumination? *i-Perception, 5*(6), 497–514.

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

Todd, J. T., & Reichel, F. D. (1989). Ordinal structure in the visual perception and cognition of smoothly curved surfaces. *Psychological Review, 96*, 643–657.

Todd, J. T., & Thaler, L. (2010). The perception of 3D shape from texture based on directional width gradients. *Journal of Vision, 10*(5):17, 1–13, doi: 10.1167/10.5.17.

Wijntjes, M. W. A., Doerschner, K., Kucukoglu, G., & Pont, S. C. (2012). Relative flattening between velvet and matte 3D shapes: evidence for similar shape-from-shading computations. *Vision Research, 12*, 1–11.