Max−Planck−Institut
für biologische Kybernetik
Spemannstraße 38 • 72076 Tübingen • Germany

Technical Report No. 081

A prior for global convexity in local shape from shading

Michael S. Langer and Heinrich H. Bülthoff

July 2000

---

[1] AG Bülthoff, E–mail: michael.langer@tuebingen.mpg.de

A prior for global convexity in local shape from shading

Michael S. Langer and Heinrich H. Bülthoff

Abstract. To solve the ill-posed problem of shape–from–shading, the visual system often relies on prior assumptions, for example, that the illumination is from above or that the viewpoint is from above. Here we demonstrate that a third prior assumption is used, namely that the surface is globally convex. Using unfamiliar randomly corrugated surfaces, we find that performance in a local shape discrimination task is significant higher when the surfaces are globally convex than when they are globally concave. The results are surprising because the qualitative global shape of the surfaces are perceptually unambiguous. The results thus generalize findings such as the hollow potato illusion (Hill & Bruce, 1994) which considered global shape perception only.

1 Introduction

When light strikes a surface, the shading pattern reflected from the surface depends on the incident light distribution, on the material of the surface, and on the 3-D shape of the surface. Inferring shape from shading is impossible since there are infinitely many shapes, lighting conditions and surface reflectances that can produce a given shading pattern (D'Zmura, 1991; Belhumeur, Kriegman, & Yuille, 1997). To resolve the ambiguity between shape and shading, the visual system relies on image information other than shading and also on prior assumptions about the scene.

One specific ambiguity in shape–from–shading occurs when a Lambertian surface is illuminated under collimated lighting and viewed under orthographic projection. The ambiguity is that the same shading pattern results if the surface is reversed in depth and illuminated from a mirror symmetric direction (Rittenhouse, 1786; Brewster, 1826). For example, a valley illuminated from the right has the same appearance as a hill illuminated from the left (see Figure 1).

The visual system can resolve this depth reversal ambiguity if other information is present in the image such as shadows (Berbaum, Bever, & Chung, 1984), occluding contours (Howard, 1983; Todd & Reichel, 1989), perspective cues, stereo, etc. The visual system also resolves the depth–reversal ambiguity by making prior assumptions, for example, that a collimated light is coming from above rather than from below (Brewster, 1826; Ramachandran, 1988) or that the viewpoint is from above rather than from below (Reichel & Todd, 1990). In this paper, we refer to prior assumptions simply as "priors."

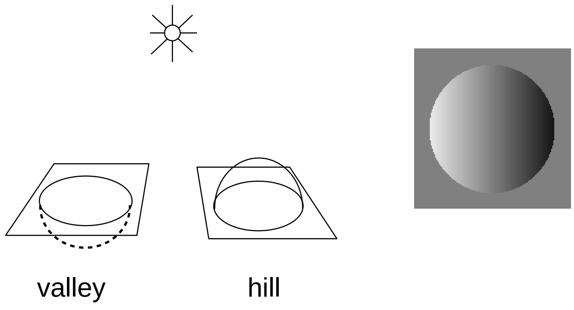


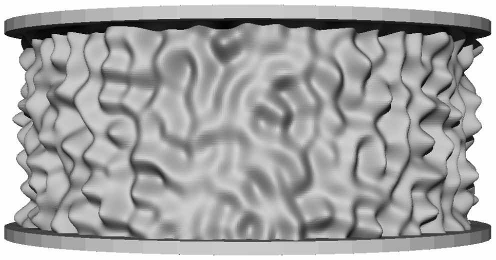
Figure 1: The depth reversal ambiguity in shading. Under collimated lighting, a hill illuminated from one direction produces the same retinal image as a valley illuminated from the opposite direction.

It has been claimed that the visual system has a prior for objects being globally convex rather than globally concave(Johnston, Hill, & Carman, 1992; Hill & Bruce, 1993). The main evidence for a prior for global convexity is that a hollow mould of an upside-down face (Hill & Bruce, 1993) or an arbitrary potato (Hill & Bruce, 1994) appears globally convex, even though the global shape of the mould is in fact concave. The "hollow potato" illusion generalizes the classical "hollow mask" illusion which applies to faces only (Luckiesh, 1916; Gregory, 1970; Yellott, l981; van den Enden & Spekreijse, 1990; Deutsch, Ramachandran, & Peli, 1990).

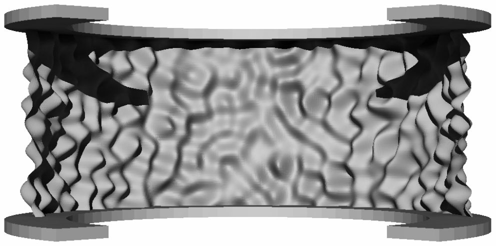
The experiments we present in this paper were motivated by two issues. First, we were concerned that the "hollow potato" illusion could have been partly due to the closed elliptical boundary of the potato. That is, the illusion of a global convex shape might not have been entirely due to a prior assumption, but rather might have depended on image information, namely the occluding contour. To address this issue, we studied perception of shape using surfaces whose global shape was readily determinable from several image cues, such as shadows, occluding contours, and perspective.

The second issue that motivated us is that, even if the visual system does have a prior for globally convex shape, it is unclear whether this prior plays a role in local shape judgments, especially on a complex corrugated surface with many local hills and valleys. For example, we might expect that if observers restricted their attention to local region of such a surface when perceiving local shape-from-shading then a prior on global convexity would play no role at all in the perceived local shape.

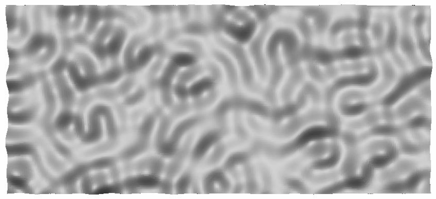
We carried out two experiments to address these issues. Both experiments used complex unfamiliar surface shapes rendered with computer graphics (see Figure 2). Each shape was either globally convex, globally concave, or globally flat. In the globally convex and concave cases, image cues were present that disambiguated the global shape, namely occluding contours, cast shadows, and perspective. The perspective cue to the global shape was particularly salient. The convex surface bulged in the middle and the concave surface was shrunken in the middle because of the differing distances of these regions from the viewer.



globally convex



globally concave



globally flat

Figure 2: An example of a globally convex, globally concave, and globally flat surface as used in Experiment 1.

The experiments were similar to Experiment 1 of (Langer & Bülthoff, 2000). We tested how well the observers could discriminate the local qualitative shape of isolated marked points on each surface. Observers judged whether each point was on a local hill or in a local valley. In (Langer & Bülthoff, 2000), we tested only the globally convex case and found that observers were well above chance. Given the depth reversal ambiguity, observers could have achieved this above-chance performance in one of two ways. They could have used the non-shading cues such as occluding contours, shadows, perspective to determine that the surface was globally convex, thereby resolving the depth reversal ambiguity. Alternatively, they could have ignored these non-shading cues and instead assumed (correctly in this case) that the surfaces were globally convex.

Which of these strategies did the visual system use? In this paper, we address this question by adding a globally concave condition to Experiment 1 of (Langer & Bülthoff, 2000). If observers use the non–shading cues to resolve the depth reversal ambiguity and don't rely on a prior for global convexity, then performance in the globally convex and globally concave conditions should now be identical, since the same cues are present in both conditions. If, on the other hand, observers ignore the non-shading cues and instead rely only on a prior for global convexity then performance should be above chance in the globally convex condition as in (Langer & Bülthoff, 2000) but below chance in the globally concave condition, and at chance overall. A combination of the two strategies is also possible.

2 Experiment 1

We present only a summary of the method. The reader is referred to Experiment 1 of (Langer & Bülthoff, 2000) for more details.

2.1 Method

2.1.1 Stimuli

Surface shapes were defined by modulating either the radius of a half-cylinder or the height of a rectangle with low pass filtered white noise (see Fig. 2). Surfaces were rendered using RADIANCE computer graphics software (Ward, 1994; Larson & Shakespeare, 1998). Surfaces were Lambertian with a reflectance of 30 percent. Interreflections were computed to two bounces.

Each surface was rendered under three collimated source conditions:

- line–of–sight, $\mathbf{L} = (0,\ 0,\ 1)$,
- above–left, $\mathbf{L} = (-.05,\ .2, 1)$,
- below–right, $\mathbf{L} = (.05,\ -.2,\ 1)$,

where $(0, 0, 1)$ is the viewing direction, and the $(x, y, 0)$ is the image plane. A weak diffuse source was added to each collimated source to simulate secondary illumination. The above–left source was used rather than a source from directly above, following the finding in (Sun & Perona, 1998) that the visual system prefers light from above–left.

Images were presented achromatically on a CRT monitor that was calibrated so that screen luminance was linearly related to rendered surface irradiance. Surfaces were presented on a uniform white background. Observers wore an eye patch over the non-dominant eye and viewed the stimuli in a dark room at a distance of 80 cm. This provided roughly the correct perspective. (Head movements were not restricted.) Each surface subtended a visual angle of roughly 20×10 degrees.

Marked points were chosen from the central 6×6 degree region. The principal curvatures of the surface at each probe were required to be either well above zero (hill condition) or well below zero (valley condition).

2.1.2 Observers

Eight observers participated (age 18–30) and were paid at a rate of 15 German Marks per hour. All observers had normal or corrected–to–normal vision.

2.1.3 Procedure

Each trial consisted of the following. A priming image was presented for 0.2 seconds. The global shape in the priming image was a flat rectangle, a concave half cylinder, or a convex half cylinder and in each case the surface was illuminated under the above–left lighting condition. A small black square probe was then superimposed on the priming image for 0.8 seconds, during which the observer made an eye movement to the probe. An image of a randomly corrugated surface such as in Fig. 2 then replaced the priming image and a reduced-size probe remained superimposed on the rendered image. In each trial, the surface in the rendered image had the same global shape as the surface in the priming image.

The observers' task was to judge whether the marked surface point was "on a hill" or "in a valley." The observers responded by pressing on one of two response keys. A three second limit was placed on the response time.

Prior to the experiment, each observer ran a practice session of about 20 trials, until they were comfortable with the task. No feedback was given either in the practice session or during the experiment. The experiment lasted roughly 15 minutes.

2.1.4 Design

A two–factor within-observer design was used with three levels per factor. The two factors were lighting direction (above–left, line–of–sight, below–right) and global shape (convex, concave, flat). Each observer ran 315 trials consisting of 35 trials for each of the nine conditions ($9 = 3 \times 3$). The order of the 315 trials was randomized for each observer.

2.2 Results and Discussion

A percent correct score was computed for each observer and for each of the nine conditions. The mean values for each condition are shown in Figure 3. An analysis of variance (ANOVA) was carried out using UNIX—STAT (Perlman, 1986).

We first considered performance in the globally flat condition to see whether observers could perform the task using shading information alone. In the globally flat condition, there were no shadow, occluding contour, or perspective cues. A one way ANOVA within the flat global shape condition revealed a main effect for light source direction ($F(2, 14) = 44.2, p < .001$). Above–left was strongly preferred over below–right as expected. Thus, shading alone was sufficient to perform the task.

We next ignored the flat global shape condition and carried out a two way ANOVA with two levels of global shape factor and three levels of light source factor. We found a main effect for global shape ($F(1, 7) = 33.0, p = .001$) with convex preferred over concave. We also found a main effect for light source direction ($F(2, 14) = 80.6, p < .001$) with light from above preferred over light from below. No interaction was found ($F(2, 14) = 1.6, p = 0.23$).

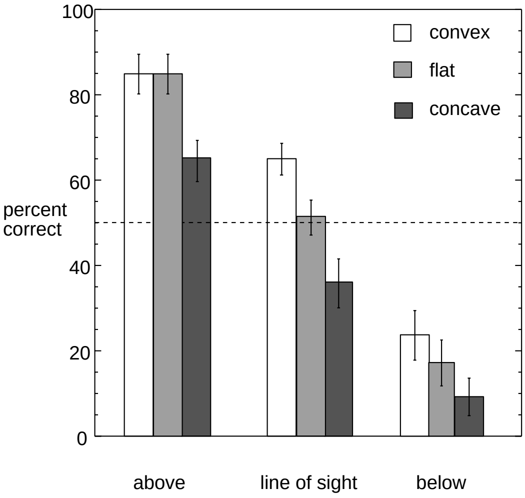


Figure 3: Mean percent correct scores over eight observers are shown for the nine combinations of global shape and light source direction. The error bars indicate the standard error of the mean.

Observers clearly relied on priors both for global convexity and for light from above. Moreover, observers did not appear to use the non–shading cues in the image, namely occluding contours, shadows, and perspective, as overall performance was near chance (47 %). The fact that observers did not use these cues was surprising. Response times were typically well within the three second limit. Mean response time was 975 ms with a standard deviation of 225 ms. Observers were thus confident enough of their local shape percepts based on shading and on their priors that they did feel the need to verify these judgments against these other image cues that were present.

3 Experiment 2

The second experiment was similar to the first, but extended the set of priors that the observer could use to resolve the depth reversal ambiguity. The third prior we allowed was viewpoint. It has been shown that when a terrain surface is viewed from an oblique angle, observers prefer a floor interpretation over a ceiling interpretation (Reichel & Todd, 1990; Mamassian & Landy, 1997, 1998). That is, they prefer an interpretation in which they are observing the surface from above rather than from below. We addressed whether this viewpoint-from-above prior is used in local shape perception, and if so, what is the strength of this prior relative to the light–from–above and globally convex shape priors.

In Experiment 1, the surfaces were oriented such that an viewpoint-from-above prior played no role. In Experiment 2, we allowed the viewpoint prior to play a role by rotating the surfaces by 90 degrees about the line of sight (see Figure 4). In the globally convex condition, the upper half of the surface had a floor orientation and the lower half of the surface had a ceiling orientation. In the globally concave condition, the opposite occurred namely the upper half of the surface had a ceiling orientation and the lower half had a floor orientation. The method was the same as in Experiment 1 apart from a few changes that we highlight below.

3.1 Method

3.1.1 Stimuli

The surfaces were rotated by 90 degrees around the line of sight prior to rendering. The globally flat condition was not included.

Probe points were chosen from regions of the cylinder that were between six and 16 degrees away from the horizontal plane, as measured from the central axis of the cylinder. These regions are marked in Figure 4.

The viewing distance was as in Experiment 1 but now a chin rest was used to restrict head movements, and thereby ensure the correct viewing perspective. The head was positioned such that the line of sight of each viewer passed through the center of the rendered image when the observer viewed the screen from the perpendicular direction. Ensuring the correct perspective was particularly important in Experiment 2 since we considered viewing direction (floor vs. ceiling) as one of the factors.

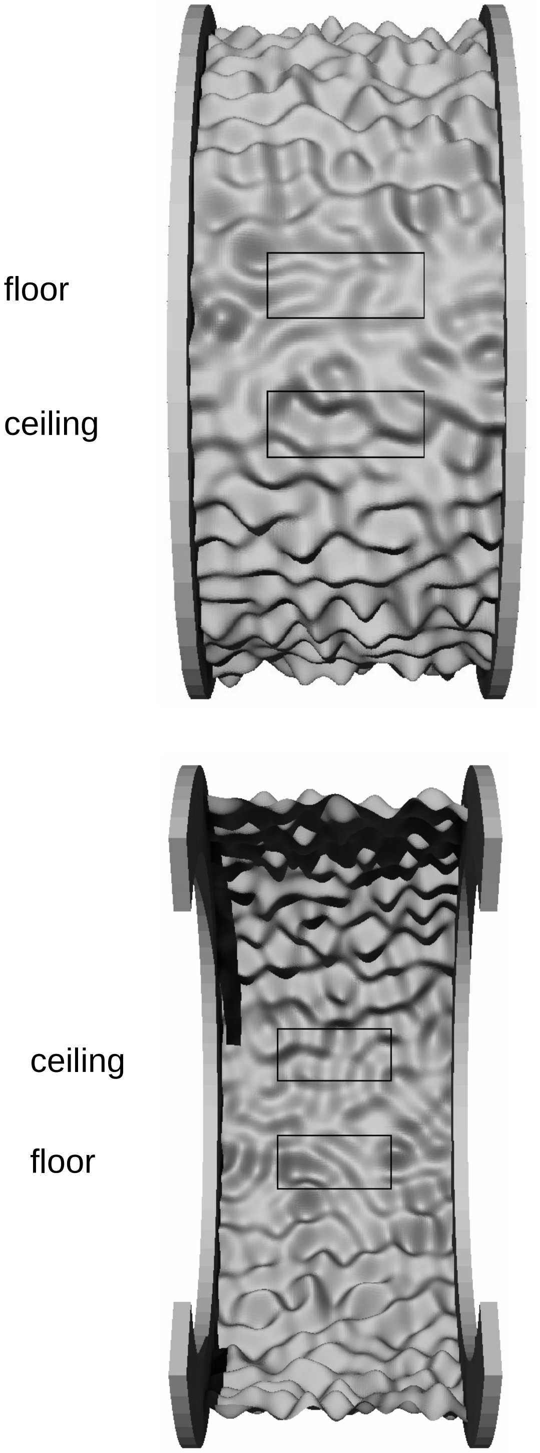


Figure 4: The rectangles show the floor and ceiling regions from which the probe points were chosen in Experiment 2.

3.1.2 Observers

Twelve new observers participated.

3.1.3 Design

A three factor within–observer design was used with two levels per factor. The three factors were light source direction (above–left, below–right), global shape (convex, concave) and viewing direction (floor, ceiling).

Observers ran 512 trials with 64 trials in each of the eight condition ($8 = 2 \times 2 \times 2$). The trials were randomly ordered for each observer. The experiment lasted roughly 30 minutes.

3.2 Results and Discussion

The results are shown in Figure 5. A three way ANOVA yielded main effects for all three factors. The strongest effect was for global shape ($F(1, 11) = 46.1, p < .001$), convex being preferred over concave. Light source from above was also preferred over light source from below ($F(1, 11) = 6.8, p = .025$) and viewpoint from above was preferred over viewpoint from below ($F(1, 11) = 9.5, p = .01$). We also found an interaction between global shape and viewpoint ($F(1, 11) = 11.6, p = .006$).

To estimate the relative strength of the three factors, we computed a linear regression of the probability of correct responses over the three factors, using the MATLAB^{TM} routine **regress**. This yielded the following fit:

$$p(L, G, V) \; = \; .51 + .1 \; L + .13 \; G + .11 \; V$$

where $L, G, V \in \{-1, 1\}$ represent the light source direction, global shape, and viewpoint variables. (The value 1 is the preferred level and the value -1 is the non–preferred level.) Thus, all three priors had roughly the same strength in our experiment.

Finally, we note that observers were again at chance overall (51 %) as they were in Experiment 1, indicating that non-shading image cues such as cast shadows, occlusion contours and perspective were not used to resolve the depth reversal ambiguity. Rather, observers resolved the ambiguity by relying entirely on prior assumptions about the scene.

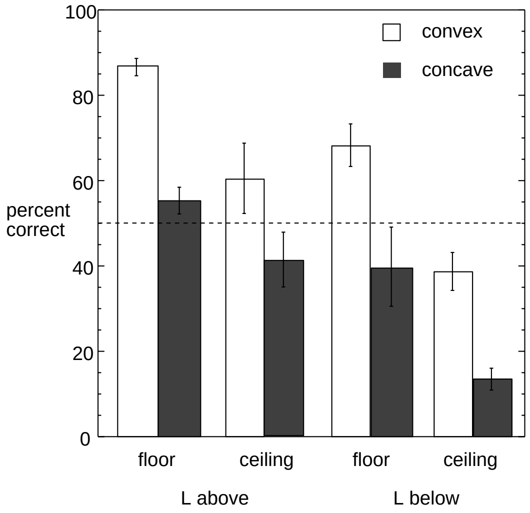


Figure 5: Percent correct scores for Experiment 2. Error bars denote the standard error of the mean for each condition over 12 observers.

4 Conclusion

We have found that a prior for globally convex shape plays a strong role in local shape-from-shading perception. For our stimuli and task, the prior on a globally convex shape had roughly the same strength as two well-known priors, namely light from above and viewpoint from above. We also found that observers did not make use of non–shading cues such as occluding contours, shadows, and perspective when performing the local shape discrimination task. Rather, they used the information in a local region of the image only – namely the local shading – and relied on prior assumptions about the scene beyond that local region.

In future work, we will address two issues that concern the spatial scale at which shading information is analyzed. For the surfaces we used in our study, it was meaningful to distinguish between global and local scales. The global scale was defined by the curvature of the half cylinder and the local scale was defined by the differential geometric curvature of the surface at the marked points. For a general surface, there may be a continuum of scales that must be considered. For example, a surface might be convex at a global scale (a solid), concave at an intermediate scale, and convex again at the local scale (a hill). It is possible that the visual system uses different priors on shape at different scales and that these priors might interact in an interesting way.

The second issue concerns the visual angle at which the surfaces are presented. The surfaces in our study all subtended a visual angle of 20 × 10 degrees. It is possible though that the prior on global shape could change qualitatively as a function of the angular size of the stimulus. For objects sub-tending a wide enough field of view, one might expect the prior on global shape to switch from a preference for global convexity to a preference for global concavity. Such a switch would be consistent with the fact that surfaces sub-tending a very large visual angle (180 degrees, say) are often the interior boundary of a closed hollow space, such as the inside of a room, rather than the closed boundary of a globally convex solid shape. That is, at large visual angles, one might find a prior for globally concave shape rather than globally convex shape, even in a local shape–from–shading task.

Acknowledgments

This research was supported by an Alexander von Humboldt Research Fellowship and a Max-Planck-Institute Stipendium to Michael Langer.

References

Belhumeur, P., Kriegman, D., & Yuille, A. (1997). The Bas-Relief Ambiguity. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 1060–1066 San Juan, Puerto Rico.

Berbaum, K., Bever, T., & Chung, C. (1984). Extending the perception of shape from known to unknown shading. *Perception*, **13**, 479–488.

Brewster, D. (1826). On the optical illusion of the conversion of cameos into intaglios and of intaglios into cameos, with an account of other analogous phenomena. *Edinburgh Journal of Science*, **4**, 99–108.

Deutsch, J. A., Ramachandran, V. S., & Peli, E. (1990). Binocular Depth Reversals Despite Familiarity Cues: An Artifact?. *Science*, **249**(4968), 565.

D'Zmura, M. (1991). Shading Ambiguity: Reflectance and Illumination. In M. S. Landy & J. A. Movshon (Eds.), *Computational Models of Visual Processing.* Cambridge, MA: MIT Press.

Gregory, R. (1970). *The Intelligent Eye.* New York: MacGraw-Hill.

Hill, H., & Bruce, V. (1993). Independent effects of lighting, orientation, and stereopsis on the hollow-face illusion. *Perception*, **22**, 887–897.

Hill, H., & Bruce, V. (1994). A comparison between the hollow-face and 'hollow-potato' illusions. *Perception*, **23**, 1335–1337.

Howard, I. (1983). Occluding edges in apparent reversal of convexity and concavity. *Perception*, **12**, 85–86.

Johnston, A., Hill, H., & Carman, N. (1992). Recognizing faces: effects of lighting direction, inversion, and brightness reversal. *Perception*, **21**, 365–375.

Langer, M. S., & Bülthoff, H. H. (2000). Depth discrimination from shading under diffuse lighting. *Perception*, **in press**.

Larson, G. W., & Shakespeare, R. (1998). *Rendering with Radiance: the art and science of lighting visualization* : Morgan Kaufmann Publishers, Inc.

Luckiesh, M. (1916). *Light and shade and their applications* (1st edition edition) : Van Nostrand.

Mamassian, P., & Landy, M. S. (1997). Illuminant and viewpoint biases from embossed surfaces. *Perception (supplement)*, **26**. ECVP abstract.

Mamassian, P., & Landy, M. S. (1998). Observer biases in the 3D interpretation of line drawings. *Vision Research*, **38**, 2817–2832.

Perlman, G. (1986). Report on UNIX—STAT release 5.1: Data analysis programs for UNIX and MSDOS. *Behavior Research Methods, Instruments & Computers*, **18**(2), 168–176.

Ramachandran, V. (1988). Perception of Shape from Shading. *Nature*, **331**(6152), 163–165.

Reichel, F. R., & Todd, J. T. (1990). Perceived Depth Inversion of Smoothly Curved Surfaces Due to Image Orientation. *Journal of Experimental Psychology: Human Perception and Performance*, **16**(3), 653–664.

Rittenhouse, D. (1786). Explanation of an optical deception. *Transactions of the American Philosophical Society*, **2**, 37–43.

Sun, J., & Perona, P. (1998). Where is the sun? *Nature Neuroscience*, **1**(3), 183–184.

Todd, J., & Reichel, F. (1989). Ordinal Structure in the Visual Perception and Cognition of Smoothly Curved Surface. *Psychological Review*, **96**(4), 643–657.

van den Enden, A., & Spekreijse, H. (1990). Binocular Depth Reversals Despite Familiarity Cues. *Science*, **249**(4968), 565–566.

Ward, G. (1994). The Radiance Lighting Simulation and Rendering System. *Computer Graphics*, 459–472.

Yellott, J. I. (l981). Binocular Depth Inversion. *Scientific American*, **245**, 118–125.