Perception, 2001, volume 30, pages 403 – 410

DOI:10.1068/p3178

A prior for global convexity in local shape-from-shading

Michael S Langer¶, Heinrich H Bülthoff
Max Planck Institute for Biological Cybernetics, Spemannstrasse 38, D 72076 Tübingen, Germany;
e-mail: michael.langer@tuebingen.mpg.de, heinrich.buelthoff@tuebingen.mpg.de
Received 22 July 2000, in revised form 30 November 2000

Abstract. To solve the ill-posed problem of shape-from-shading, the visual system often relies on prior assumptions such as illumination from above or viewpoint from above. Here we demonstrate that a third prior assumption is used—namely that the surface is globally convex. We use complex surface shapes that are realistically rendered with computer graphics, and we find that performance in a local-shape-discrimination task is significantly higher when the shapes are globally convex than when they are globally concave. The results are surprising because the qualitative global shapes of the surfaces are perceptually unambiguous. The results generalise findings such as the hollow-potato illusion (Hill and Bruce 1994 *Perception* **23** 1335 – 1337) which consider 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 three-dimensional (3-D) shape of the surface. Inferring exact 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 et al 1997). To resolve these ambiguities between shape and shading, the visual system relies on image information in addition to shading. It also relies 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

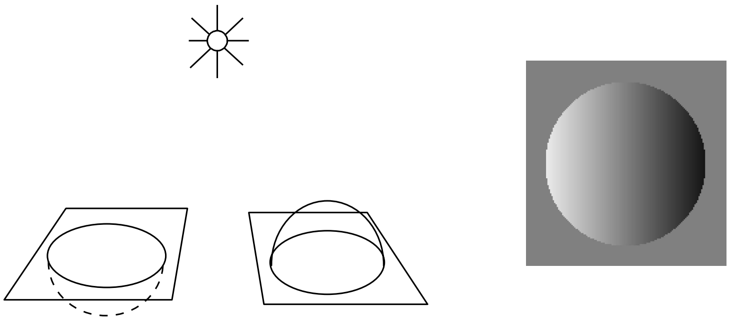


Figure 1. An example of the depth-reversal ambiguity in shading. Under collimated lighting such as on a sunny day, a valley illuminated from the right produces the same retinal image as a hill illuminated from the left.

¶ Correspondence and requests for reprints should be addressed to: Michael Langer, School of Computer Science, McGill University, Montréal, Québec H3A 2A7, Canada.

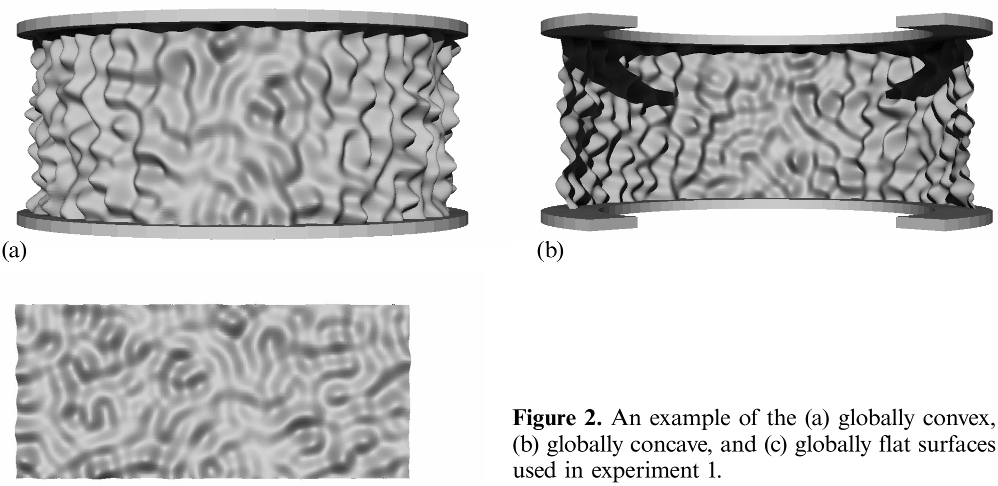
ambiguity if other information is present in the image, such as shadows (Berbaum 1984), occluding contours (Howard 1983; Todd and Reichel 1989), perspective cues, stereo, etc. The visual system can also resolve the depth-reversal ambiguity by making prior assumptions about the scene, for example that the illumination is from above rather than from below (Brewster 1826; Ramachandran 1988), or that the viewpoint is from above rather than from below (Reichel and Todd 1990). In this paper, we refer to such assumptions as 'priors'.

It has been claimed that the visual system uses a prior that objects are globally convex rather than globally concave (Johnston et al 1992; Hill and Bruce 1993). The main evidence for such a prior is that a hollow mould of an upside-down face (Hill and Bruce 1993) or of an arbitrary potato (Hill and Bruce 1994) appears globally convex, even though the global shape of the mould is in fact concave. The latter hollow-potato illusion generalises the classical hollow-mask illusion which applies to faces only (Luckiesh 1916; Gregory 1970; Yellott 1981; Deutsch et al 1990; van den Enden and Spekreijse 1990).

The experiments we present in this paper were motivated by two issues related to the hollow-potato illusion. First, we were concerned that the hollow-potato illusion might not have been due to a prior for global convexity but rather may have been due to the near-elliptical boundary contour of the potato which may have suggested a globally convex shape. To address this possible confound, we decided to study more complex surfaces for which several types of shape information were present, and also to consider both convex and concave global shapes.

The second issue that motivated us was that, even if the visual system does have a prior for global convexity, it is unclear whether this prior would play a role in local-shape judgments. For a task in which the observers restrict their attention to a local region of a surface only, one might expect a prior on global convexity to play no role.

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 globally concave cases, several image cues were present that disambiguated the global shape, such as 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.



(a) (b) (c)

Figure 2. An example of the (a) globally convex, (b) globally concave, and (c) globally flat surfaces used in experiment 1.

The experiments were similar to one of our previous experiments [experiment 1 of Langer and Bülthoff (2000)]. We tested how well the observers could discriminate the local qualitative shape of isolated points on a surface. Observers judged whether a point was on a local hill or in a local valley. Previously, we tested only the globally convex case and found that observers were well above chance (Langer and Bülthoff 2000). In light of the depth-reversal ambiguity, we know that observers could have achieved this above-chance performance in one of two ways. They could have used the non-shading cues (occluding contours, shadows, perspective) to determine that the surfaces were globally convex, and thereby resolve the depth-reversal ambiguity. That is, the depth-reversal ambiguity applies to the whole surface, and so, if one can resolve the ambiguity globally, then one can also resolve it locally. The second strategy is that they could have ignored the non-shading cues and instead assumed a priori that the surfaces were globally convex, again resolving the depth-reversal ambiguity both globally and locally.

Which of these strategies does the visual system use? To address this question in the present paper, we repeat our previous experiment, and add a globally concave condition. If observers use the non-shading cues to resolve the depth-reversal ambiguity rather than relying on a prior, then performance in the globally convex and globally concave conditions should be identical, since roughly the same cues are present in both conditions. If, on the other hand, observers ignore the non-shading cues and instead rely on a prior then performance should be above chance in the globally convex condition (as in Langer and Bülthoff 2000) but below chance in the globally concave condition, and at chance overall.

2 Experiment 1

We present only a summary of the method. The reader is referred to experiment 1 of Langer and 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 figure 2). Surfaces were rendered with the use of RADIANCE computer graphics software (Ward 1994; Ward Larson and Shakespeare 1998). Surfaces were Lambertian with a reflectance of 30%. Interreflections were computed to two bounces.

Each surface was rendered under three collimated source conditions:

- line-of-sight, light source (0, 0, 1)
- above-left, light source (−0.05, 0.2, 1)
- below-right, light source (0.05, −0.2, 1)

where (0, 0, 1) is the viewing direction. 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 and 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 nondominant eye and viewed the stimuli in a dark room at a distance of 80 cm. Each surface subtended a visual angle of roughly 20 deg × 10 deg, which provided roughly the correct perspective. (Head movements were not restricted.)

Marked points were chosen from the central 6 deg × 6 deg region. The principal curvatures of the surface at each probe were required to be well above zero (hill condition) or well below zero (valley condition). Further details on the criteria of probe points are given in Langer and Bülthoff (2000). Although the probes were chosen randomly and automatically, we did verify that each probe used was correctly placed on a hill or valley. There were a few cases of a 'crater on top of a volcano' but we decided to leave such cases in, rather than to remove them ad hoc by hand.

2.1.2 *Observers.* Eight observers participated (aged 18–30 years) and were paid at a rate of DM 15 per hour. All observers had normal or corrected-to-normal vision.

2.1.3 *Procedure.* Each trial consisted of the following. First, a priming image was presented for 0.2 s. The global shape in each priming image was either a flat rectangle, a concave half-cylinder, or a convex half-cylinder. In each of these priming images, the surface was illuminated under the above-left lighting condition. This is the preferred lighting condition for shape-from-shading perception (Sun and Perona 1998). We used this lighting condition to prime the observer about the global shape of the test surface to follow.

After this initial priming image, a small black square probe was superimposed on the priming image for 0.8 s, during which time the observer made an eye movement to the probe. An image of a randomly corrugated surface such as in figure 2 then replaced the priming image and a reduced-size probe remained superimposed on the rendered image. The surface in the rendered image had the same global shape as the surface in the priming image.

On each trial the task was to judge whether the marked point was "on a hill" or "in a valley". Observers responded by pressing one of two response keys. A 3 s limit was placed on the response time.

Prior to the experiment, each observer ran a practice session of 20 trials. No feedback was given either in the practice session or during the experiment. The experiment lasted roughly 15 min.

One minor point should be discussed, concerning why we used overhead lighting in each priming stimulus. If we had, instead, used the lighting condition which corresponds to the test stimulus, then, in the condition of a globally concave surface and under lighting from below, the observers would *misperceive* the global shape of the priming surface, because of a prior that the light source is from above. Since the observer would misperceive the global shape of the priming stimulus in this condition, he or she would be incorrectly primed for the global shape of the test stimulus. We did not want to put the observers at such a disadvantage. Rather, our goal in priming them was to provide information about the upcoming global shape, so that they could rely less on their prior for global shape (if, indeed, there is such a prior).

2.1.4 *Design.* A two-factor within-observers 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 randomised for each observer.

2.2 *Results and discussion*

A percentage-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 with UNIX–STAT (Perlman 1986).

We first consider the globally flat condition. In this condition, the only cue to shape other than shading is a small perspective distortion near the boundary of the surface. We regard this cue as perceptually insignificant for the task because the probe points were far away from this boundary. Thus, the globally flat condition tests how well observers perceive local shape-from-shading using shading information alone.

A one-way ANOVA within the globally flat condition revealed a main effect for light-source direction ($F_{2,14} = 44.2$, $p < 0.001$). Above-left was strongly preferred over below-right, as expected. Thus, shading information alone was sufficient to perform

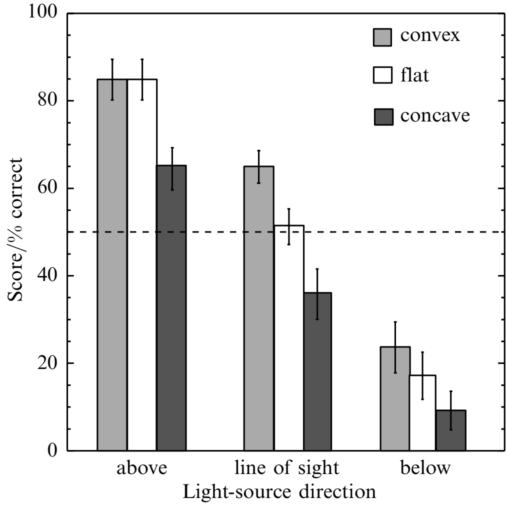


Figure 3. Mean percentage-correct scores over the eight observers in experiment 1 are shown for the nine combinations of global shape and light-source direction. Error bars indicate the standard error of the mean.

the task. Performance when the light source was at the line of sight was at chance (52%) which is not surprising, since there is no way for the visual system to resolve the depth-reversal ambiguity in this condition.

We next ignored the flat-global-shape condition and carried out a two-way ANOVA with only two levels of global-shape factor and with three levels of light-source factor. We found a main effect for global shape ($F_{1,7} = 33.0$, $p = 0.001$) with convex preferred over concave. We also found a main effect for light-source direction ($F_{2,14} = 80.6$, $p < 0.001$) with light from above preferred over light from below. No interaction was found ($F_{2,14} = 1.6$, $p = 0.23$).

Observers clearly relied on priors both for global convexity and for light from above. Moreover, observers did not appear to use the non-shading information in the image as overall performance was near chance (47%). The fact that observers did not use the non-shading cues was surprising. Response times were typically well within the 3 s 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 priors that they did not need to validate these judgments with the non-shading cues.

In summary, we believe that observers resolve the depth-reversal ambiguity in shape-from-shading using priors, both on global shape and on light-source direction. When the priors are incorrect, as is the case when the surface is globally concave, the visual system tends to misperceive the sign of curvature of the local shape. This is a direct result of perceiving the depth-inverted surface rather than the correct surface.

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 considered 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 and Todd 1990; Mamassian and 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 the strength of this prior is relative to the light-from-above and global convexity priors seen in experiment 1.

In experiment 1, the surfaces were oriented such that a 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° about the line of sight (see figure 4). In the globally convex

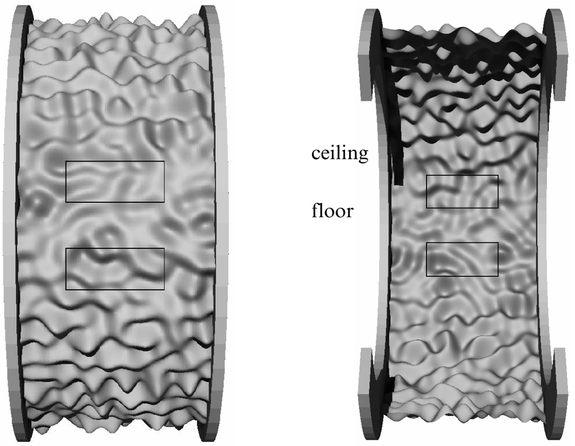


Figure 4. The rectangles show the floor region (viewpoint from above) and ceiling region (viewpoint from below) from which the probe points were chosen for (a) convex and (b) concave surfaces used in experiment 2.

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° 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 6° and 16° 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 centre of the rendered image when the observer viewed the screen from the perpendicular direction. Ensuring the correct perspective was important in experiment 2 since we considered viewing direction (floor versus ceiling) as one of the factors.

3.1.2 *Observers.* Twelve new observers participated.

3.1.3 *Design.* A three-factor within-observers 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 conditions ($8 = 2 \times 2 \times 2$). The trials were randomly ordered for each observer. The experiment lasted roughly 30 min.

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 ($F_{1,11} = 46.1$, $p < 0.001$), convex being preferred over concave. Light source from above was also preferred over light source from below ($F_{1,11} = 6.8$, $p = 0.025$), and viewpoint from above was preferred over viewpoint from below ($F_{1,11} = 9.5$, $p = 0.01$). An interaction was found between global shape and viewpoint ($F_{1,11} = 11.6$, $p = 0.006$).

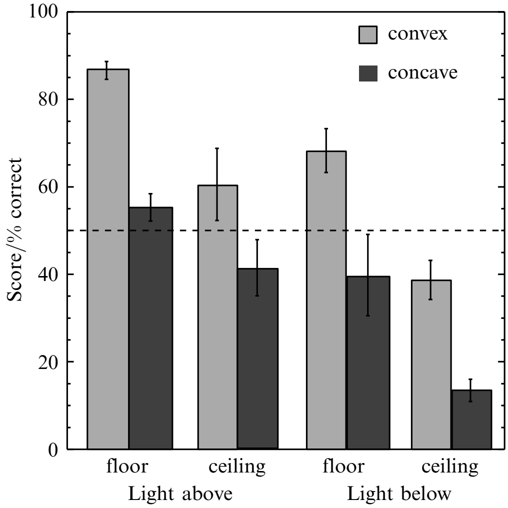


Figure 5. Mean percentage-correct scores for experiment 2 over the twelve observers. Error bars indicate the standard error of the mean.

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™ routine regress. This yielded the following fit:

$$p(L,\ G,\ V) = 0.51 + 0.1L + 0.13G + 0.11V$$

where $L,\ G,\ V \in \{-1,\ 1\}$ represent the light-source direction, global shape, and viewpoint variables, respectively. (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.

Note that observers were again at chance overall (51%), as they were in experiment 1, indicating that non-shading cues (shadows, occlusion contours, perspective) were not used to resolve the depth-reversal ambiguity. Rather, observers resolved the ambiguity by relying on prior assumptions about the scene.

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: light from above and viewpoint from above. We also found that observers did not use image cues such as occluding contours, shadows, and perspective to perform the local-shape-discrimination task. Rather, they used the information in a local region of the image only: the local shading.

In future work, we plan to address two issues that concern the spatial scale at which shading information is analysed. 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, concave at an intermediate scale, and convex again at the local scale. 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.

A second issue concerns the visual angle at which the surfaces are presented. The surfaces in our study all subtended a visual angle of 20 deg × 10 deg. It is possible, however, that the prior on global shape could change qualitatively as a function of the angular size of the stimulus. For objects subtending 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 a surface subtending a very large visual angle (180 deg, say) is often the interior boundary of a closed hollow space such as the inside of a room, rather than the closed boundary of a convex solid shape. It would be interesting if, at large visual angles, one were to find a prior for globally concave shape rather than a prior for globally convex shape, even in a local shape-from-shading task.

Acknowledgments. Michael Langer was supported by an Alexander von Humboldt Research Fellowship and by a stipend from the Max Planck Society.

References

Belhumeur P, Kriegman D, Yuille A, 1997 "The bas-relief ambiguity", in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, San Juan, Puerto Rico, June 1997* (Washington, DC: IEEE Computer Society Press) pp 1060 – 1066

Berbaum K, Bever T, Chung C S, 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** 565

D'Zmura M, 1991 "Shading ambiguity: Reflectance and illumination", in *Computational Models of Visual Processing* Eds M S Landy, J A Movshon (Cambridge, MA: MIT Press) pp 187 – 207

Enden A van den, Spekreijse H, 1990 "Binocular depth reversals despite familiarity cues" *Science* **249** 565 – 566

Gregory R, 1970 *The Intelligent Eye* (New York: McGraw-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 "Recognising 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* **29** 649 – 660

Luckiesh M, 1916 *Light and Shade and their Applications* first edition (New York: Van Nostrand)

Mamassian P, Landy M S, 1997 "Illuminant and viewpoint biases from embossed surfaces" *Perception* **26** Supplement, 51

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** 168 – 176

Ramachandran V, 1988 "Perception of shape from shading" *Nature* **331** 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** 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** 183 – 184

Todd J, Reichel F, 1989 "Ordinal structure in the visual perception and cognition of smoothly curved surface" *Psychological Review* **96** 643 – 657

Ward G, 1994 "The radiance lighting simulation and rendering system" *Computer Graphics* 459 – 472

Ward Larson G, Shakespeare R, 1998 *Rendering with Radiance: The Art and Science of Lighting Visualization* (San Francisco, CA: Morgan Kaufmann)

Yellott J I, 1981 "Binocular depth inversion" *Scientific American* **245**(1) 118 – 125