

Shape from shading under inconsistent illumination

John D. Wilder

Department of Psychology and Centre for Vision
Research, York University, Toronto, Canada
Department of Psychology, University of Toronto,
Toronto, Canada



Wendy J. Adams

School of Psychology, University of Southampton,
Southampton, UK



Richard F. Murray

Department of Psychology and Centre for Vision
Research, York University, Toronto, Canada



People are able to perceive the 3D shape of illuminated surfaces using image shading cues. Theories about how we accomplish this often assume that the human visual system estimates a single lighting direction and interprets shading cues in accord with that estimate. In natural scenes, however, lighting can be much more complex than this, with multiple nearby light sources. Here we show that the human visual system can successfully judge 3D surface shape even when the lighting direction varies from place to place over a surface, provided the scale at which these lighting changes occur is similar to, or larger than, the size of the shape features being judged. Furthermore, we show that despite being able to accommodate rapid changes in lighting direction when judging shape, observers are generally unable to detect these changes. We conclude that, rather than relying on a single estimated illumination direction, the human visual system can accommodate illumination that varies substantially and rapidly across a surface.

Introduction

Shape from shading relies on the relationship between surface attitude and the illumination direction: Matte surfaces reflect more incident illumination if they face the illuminant than if they are slanted away from it. For a perfectly Lambertian surface, surface luminance is proportional to the cosine of the angle between the illumination direction and the surface normal. To make use of this relationship between luminance and surface attitude, models of shape from shading generally assume a single known or estimated illumination direction.

In contrast, however, lighting in everyday environments is complex. It can include primary light sources such as the sun or electric lights, as well as reflective surfaces and light-scattering volumes that function as secondary light sources. The lighting conditions at any point in space therefore depend on both primary sources and on the geometry and optical properties of surrounding materials (Gershun, 1939; Moon & Spencer, 1981). Furthermore, lighting conditions generally vary across spatial locations within a scene, sometimes gradually as when moving through a large, evenly lit space, and sometimes suddenly as when crossing a shadow boundary. Fine details of lighting conditions are unimportant for the appearance of convex matte objects (Ramamoorthi & Hanrahan, 2001; Basri & Jacobs, 2003), but even coarse lighting properties, such as direction and diffuseness, can vary substantially within a scene (Dror, Willsky, & Adelson, 2004; Mury, Pont, & Koenderink, 2007, 2009a, 2009b; Cuttle, 2008; Morgenstern, Geisler, & Murray, 2014). Complex lighting poses a challenge for the human visual system, because the retinal image is generated by interactions between light, shape, and material properties, and as a result the information available at the retina is deeply ambiguous (e.g., Belhumeur, Kriegman, & Yuille, 1999). Current models of human visual perception make a wide range of claims about how we represent lighting. Some suggest that human vision represents both lighting direction and diffuseness (Boyaci, Maloney, & Hersh, 2003; Bloj et al., 2004; Morgenstern et al., 2014), while others claim that we do not represent lighting at all for some purposes, including perception of shape from shading (Fleming, Holtmann-Rice, & Bülthoff, 2011), and still others have intermediate views (Gilchrist et al., 1999). Human perception of shape from shading has been studied

Citation: Wilder, J. D., Adams, W. J., & Murray, R. F. (2019). Shape from shading under inconsistent illumination. *Journal of Vision*, 19(6):2, 1–15, <https://doi.org/10.1167/19.6.2>.

<https://doi.org/10.1167/19.6.2>

Received September 7, 2018; published June 5, 2019

ISSN 1534-7362 Copyright 2019 The Authors



extensively under collimated lighting (Todd & Mingolla, 1983; Ramachandran, 1988), and to a more limited extent under completely diffuse lighting (Langer & Bühlhoff, 2000), but much less is known about how human vision deals with the complex lighting conditions that are typical of natural scenes, including spatiotemporal lighting changes within a scene. Similarly, classic computer vision approaches to shape from shading assume a distant point light source (Horn, 1975; Zhang, Tsai, Cryer, & Shah, 1999), although recent work has made progress on relaxing this assumption (e.g., Tian, Tsui, Yeung, & Ma, 1999; Barron & Malik, 2015).

Even though lighting is complex and retinal images are ambiguous, people are nevertheless able to estimate lighting conditions in 3D pictorial space surprisingly well. Koenderink, Pont, van Doorn, Kappers, and Todd (2007) introduced a gauge object method in which observers adjust the intensity, direction, and diffuseness of the simulated illumination on a spherical probe in a complex scene so that the lighting on the probe is perceived to be correct for its location. They found that observers gave close to veridical settings, and several subsequent studies used similar methods to examine lighting perception in greater detail. Kartashova, Sekulovski, de Ridder, te Pas, and Pont (2016) used multiple probe settings throughout a scene to reconstruct a full map of perceived illumination. They found that observers generated light fields that were very similar to the true physical light field, except that the perceived light field was somewhat simplified, and biased towards diverging light fields (e.g., the pattern of light radiating outwards from a candle). Using similar methods, Xia, Pont, and Keynderickx (2016) found that the accuracy of probe matches depended on the content of the scene, and that matches were more accurate in scenes that contained objects with many planar surfaces. Furthermore, te Pas, Pont, Dalmaijer, and Hooge (2017) found that lighting percepts depended more strongly on specular highlights and shadows than on shading gradients.

Nevertheless, human vision can also be insensitive to substantial lighting changes within a scene. Ostrovsky, Cavanagh, and Sinha (2005) showed that people are poor at rapidly detecting even large changes in lighting from one place to another in some types of complex scenes. In their experiments, subjects could not detect strong, artificially created lighting inconsistencies in photographs or paintings, and had to search slowly and serially through a grid of objects to decide which one was illuminated from a different direction than the others. Ostrovsky et al. concluded that the human visual system does not constrain its estimates of lighting conditions to be consistent throughout a scene. Ostrovsky et al. did not measure performance on shape or reflectance tasks, but their stimuli had a natural

appearance despite highly inconsistent lighting. If inconsistent lighting strongly disrupted shape or reflectance estimation, then we might expect their subjects to have been able to detect such inconsistencies.

There is also evidence in studies of shape from shading that vision can accommodate spatial variations in lighting, but only variations that are typical in real scenes. Morgenstern, Murray, and Harris (2011) found that observers use lighting cues on one object to infer shape from shading on a separate object that is several degrees of visual angle away, reflecting a preference for smooth changes in lighting. van Doorn, Koenderink, and Wagemans (2011) and van Doorn, Koenderink, Todd, and Wagemans (2012) found that observers are strongly biased to perceive lighting conditions that are either collimated, like light on a sunny day, or diverging, like light radiating outwards from a candle. Observers were unable to see shaded images as illuminated by other, less common lighting patterns, such as cyclical light fields where lighting directions form a closed loop.

Presumably the perception of shape from shading breaks down if lighting varies too rapidly from place to place, but where does this breakpoint occur? An answer to this question would help us to understand the limits of the human visual system's representation of lighting. To investigate, we measured human performance in a simple shape judgment task, using computer-rendered images of bumpy surfaces under lighting conditions that varied from place to place at different rates. To preview our key findings across four experiments, (a) Shape perception was surprisingly robust to illumination manipulations: Performance was well above chance as long as lighting was roughly constant over the scale of the small bumpy features whose shape was being judged. Performance decayed when lighting changed too rapidly from place to place, at finer scales than the shape modulations. (b) Interestingly, although observers could accommodate abrupt changes in lighting direction when estimating shape, they were mostly unable to detect these lighting changes. We discuss the implications of our findings for theories of perceptual representation of lighting and perception of shape from shading.

Experiment 1

In Experiment 1 we investigated how spatial variations in lighting direction affect people's ability to perceive shape from shading. We varied lighting conditions from place to place over a bumpy surface, and examined how the rate of change affected performance on a depth judgment task.

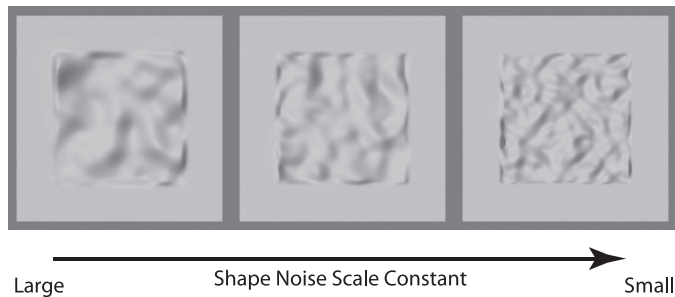


Figure 1. Examples of single lighting direction stimuli shown in Experiment 1. Here the point light source is to the left. The surfaces differ in the spatial frequency of the 2D low-pass Gaussian noise used to construct the bumpy inner regions.

Methods

Observers

There were eight observers (three male, five female; mean age 24 years). One was author JW, and the remaining observers were unaware of the purpose of the experiment and were paid for their participation. One observer's data were excluded from analysis because performance was below 60% correct in all conditions. All observers in all experiments reported normal or corrected-to-normal vision, and gave written informed consent before participating. All experimental procedures were approved by the Office of Research Ethics at York University.

Stimuli

The stimuli were computer-generated images of a square panel, rendered in RADIANCE (Ward Larson & Shakespeare, 2004) and displayed on an LCD monitor at a viewing distance of 57 cm (Figure 1). At this viewing distance, 1 cm on the virtual object subtended 1° of visual angle (degree). The stimulus was 16 cm square, including 3 cm wide flat outer borders, and a 10 cm square, bumpy inner region. The entire panel had a simulated reflectance of 30%. The inner bumpy region of each stimulus was constructed using a sample of 2D low-pass Gaussian noise, which we call the *depth map*. The value of the depth map at each point determined how far the corresponding surface point was displaced from the plane containing the flat borders, either towards or away from the observer. The depth map was created by convolving a sample of Gaussian white noise with a Gaussian convolution kernel with scale constant $\sigma_S = 0.35, 0.53, \text{ or } 0.70$ cm. Larger values of the scale constant produced wider bumps. We refer to the resulting depth maps as high, medium, and low frequency, respectively. Each depth map had a pointwise *SD* of $0.0075 \sigma_S$. Thus in a low frequency depth map, for example, a relatively large bump with a peak depth of 2 SD had a height of 0.0075

$\times 2 \sigma_S = 0.0075 \times 2 \times 0.70 \text{ cm} = 1.05 \text{ cm}$. To create a smooth transition from the inner bumpy region to the smooth borders, we multiplied the shape noise by a sigmoidal function that varied smoothly but rapidly from a value of one in the central square to a value of zero in the borders.

To create *single lighting direction* stimuli (Figure 1), we rendered the panels described above under a single distant point-like light source along with omnidirectional ambient lighting. The point-like source was an infinitely distant disk that subtended 1° and had a luminance of $1.0 \times 10^6 \text{ cd/m}^2$. The ambient illumination had a luminance of 13.5 cd/m^2 in all directions. The point source was behind the simulated camera, 30° off the line of sight (i.e., the lighting direction had a slant of 30°). With slant held constant, we varied the tilt of the point source direction from rightwards (0°), through upwards (90°), to leftwards (180°). Thus, in a Cartesian coordinate system where the camera is on the $+z$ axis looking at the origin, and the x axis is to the right and the y axis is upwards, the lighting directions were $(\sin 30^\circ \cos \theta, \sin 30^\circ \sin \theta, \cos 30^\circ)$, with $0^\circ \leq \theta \leq 180^\circ$. We rendered each panel under 181 lighting directions, with θ ranging from 0° to 180° in 1° steps, always with the ambient light source present as well. These lighting conditions and the uniform reflectance of 30% gave the flat outer panels a luminance of 24 cd/m^2 in the rendered image (4 cd/m^2 from ambient lighting, 20 cd/m^2 from the point source), and gave a surface patch directly facing the point light source a luminance of 27 cd/m^2 .

To create *lighting noise* stimuli, in which the lighting direction varied across the surface (Figure 2), we made composite images by smoothly merging many single lighting direction stimuli that shared a common depth map. A sample of 2D low-pass uniform noise defined the *lighting map*. The values in the lighting map determined the lighting direction at each point in the composite image. For example, if the top-left element of the lighting map had a value of 0, then the top-left pixel of the composite image was taken from the top-left pixel of the single lighting direction image with a lighting direction of 0° . And more generally, if element (i, j) of the lighting map had a value (rounded to the nearest integer) of k , then pixel (i, j) in the composite image was taken from pixel (i, j) in the single lighting direction image that had lighting direction k° . In this way, we merged many single lighting direction images to produce a composite image in which the local lighting direction varied smoothly from place to place.

We created lighting maps as follows: We convolved a sample of 2D Gaussian white noise (mean zero, *SD* 1) with a unit-energy Gaussian kernel with a scale constant of $\sigma_L = 0.175, 0.350, 0.525, 0.700, 1.40, \text{ or } 2.80$ cm. We passed each point of the resulting low-pass noise through the standard normal cumulative distri-

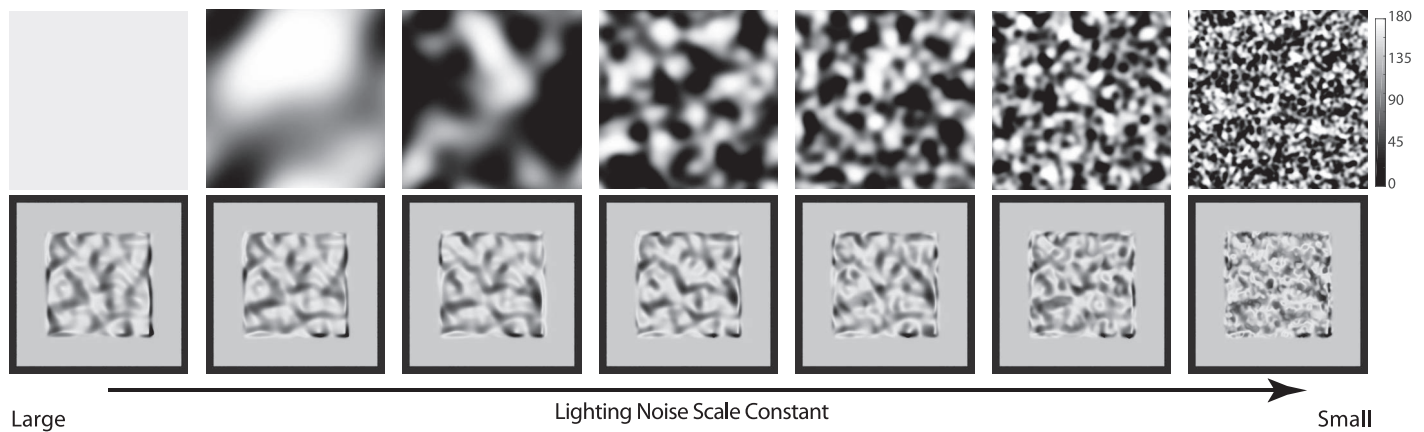


Figure 2. Examples of lighting noise stimuli shown in Experiment 1. Here a single shape is shown under several lighting conditions. The lighting direction maps, shown in the top row, are samples of low-pass 2D uniform noise. The corresponding stimulus images are shown in the bottom row. In the leftmost stimulus the lighting direction is constant, and in each stimulus to the right the lighting noise is generated with a progressively smaller scale constant, and so lighting direction varies more rapidly from place to place.

bution function. This gave a sample of low-pass noise where the marginal distribution of each element was the uniform distribution on the interval (0, 1). Finally, we multiplied this low-pass uniform noise sample by 180° , to produce a lighting map with values in the range (0° , 180°). The scale constant of the low-pass Gaussian noise, σ_L , determined the rate at which the lighting direction varied from place to place in the composite stimuli: Large scale constants produced low-frequency lighting maps, and small scale constants produced high-frequency lighting maps. Figure 3 shows how this manipulation affects the appearance of a simple and familiar object, namely a matte sphere. Thus, each point on the surface is lit by collimated light, and the direction of the light source varies across the surface.

The diffuse light source is weak enough that the dark-is-deep cue (Langer & Bühlhoff, 2000) is an unreliable cue to the surface shape, and the observers must rely on the collimated light source. For a surface only lit by the diffuse light source, the peaks have a mean luminance of 3.9 cd/m^2 , while the valleys have a

mean luminance of 3.14 cd/m^2 (a difference of 0.8 cd/m^2). For the stimuli used in the experiment, the luminance varies over a much wider range, between 14.5 cd/m^2 and 25.5 cd/m^2 . Thus, it would be difficult for observers to detect and use the dark-is-deep signal.

Observers viewed the stimuli on an LCD monitor at a viewing distance of 57 cm, with head position stabilized by a chin rest. Stimuli were shown on a gray background of luminance 6.0 cd/m^2 . The monitor had a resolution of 1920×1080 pixels, a pixel size of 0.247 mm , and a nominal frame rate of 60 Hz. We characterized the monitor's gamma function using a Minolta LS-110 photometer, and the stimulus display software inverted a fitted gamma function to show the required luminances.

Procedure

Each observer completed eight sessions of 525 trials, each of which lasted approximately 20 minutes, including a 2–3 minute break. There were 21 randomly

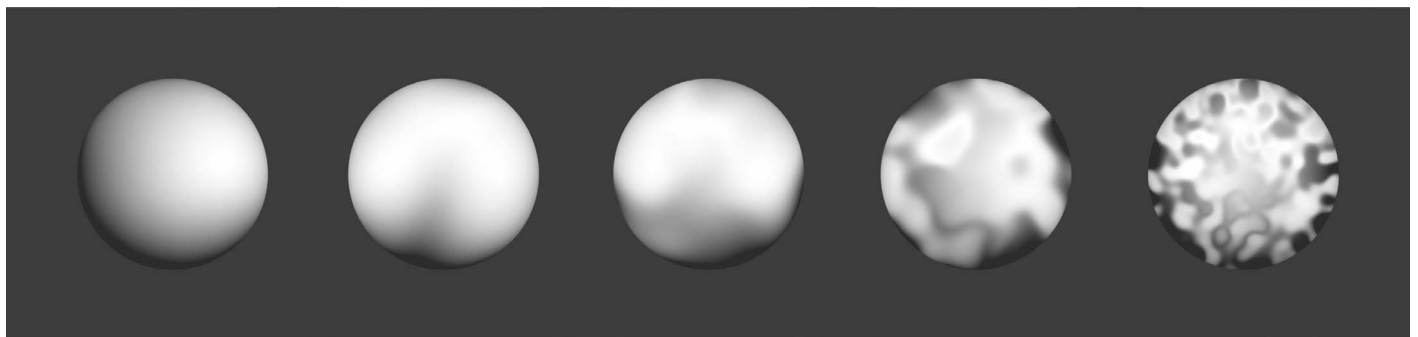


Figure 3. Examples of spheres under lighting noise. The uniformly illuminated sphere (on the left) yields a clear shape percept. The shape percept is slightly perturbed under small amounts of noise, and is almost completely destroyed under the highest frequency lighting variations (on the right).

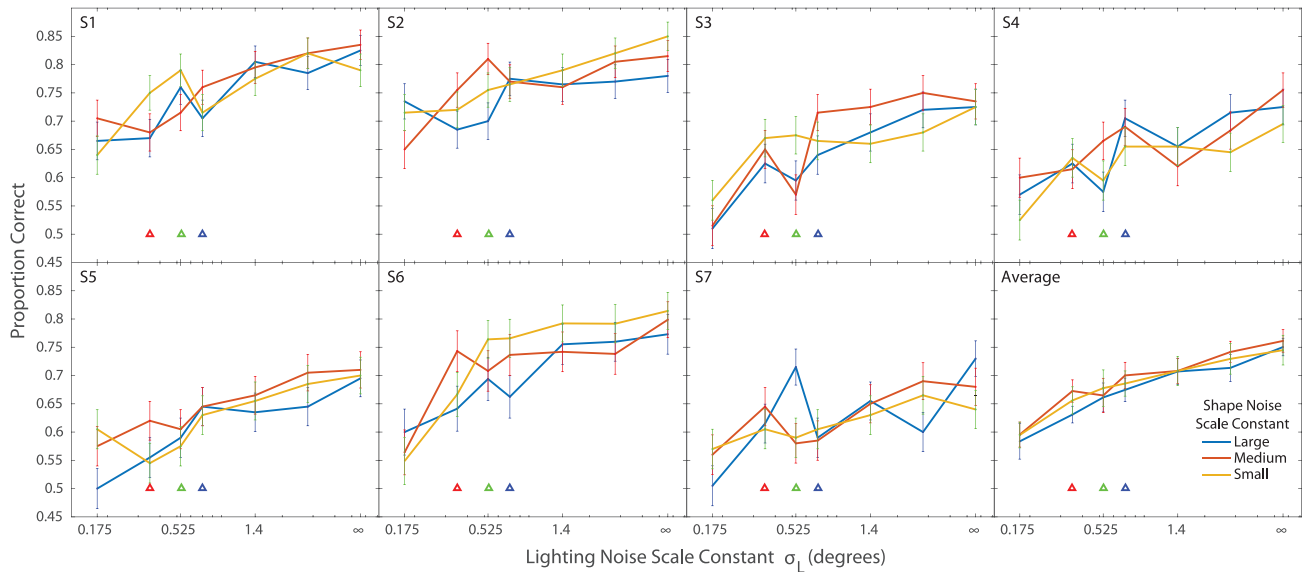


Figure 4. Results from Experiment 1. Each panel plots proportion correct in the depth discrimination task versus the lighting noise scale constant. Error bars show 1 SE.

interleaved stimulus conditions: three shape scales crossed with seven lighting scales (the six values listed in the Stimuli section, plus the single lighting direction condition, which effectively has a lighting scale constant of infinity). On each trial the observer saw a square panel with two red probe points in the inner bumpy region, separated by between 0.5° and 1.0° . The observer pressed one of two buttons to indicate whether the left or right probe was closer. The probe point locations were chosen randomly, with the constraints that (a) the points they marked on the bumpy surface had a depth difference of between 0.6 and 0.8 times the standard deviation of the depth noise, and (b) there was no change in the sign of the depth gradient along the straight line connecting the two points, i.e., they could not be separated by a peak or a trough. Pilot trials with single lighting direction stimuli showed that these criteria made the task difficult but feasible. The stimulus remained onscreen until the observer responded. No feedback was given as to whether the response was correct. After the observer responded, there was a short pause and then the next trial began. The median response time across all observers and conditions was 0.8 s.

Observers viewed 40 unique images in each of the 21 stimulus conditions. Each unique image was generated using depth and lighting maps created from independent noise samples. Each unique image was presented on five separate trials, with the red probe dots at a different location each time. This gave $40 \text{ stimuli} \times 5 \text{ repetitions} = 200 \text{ trials per condition}$, and $21 \text{ conditions} \times 200 \text{ trials} = 4,200 \text{ trials}$ for each observer in the full experiment. The 4,200 trials were divided randomly across the eight sessions.

Results and discussion

Figure 4 shows proportion correct for the shape task, as a function of the lighting scale σ_L , for individual observers and averaged across observers. Data points towards the left of each panel show performance for lighting maps at small scales, where the lighting direction varied rapidly from place to place, and data points at the far right ($\sigma_L = \infty$) show performance when illumination direction was constant across the stimulus. Performance was clearly worse in conditions where the lighting direction varied rapidly.

To see whether proportion correct was predicted better by the lighting scale in degrees of visual angle (*absolute* lighting scale), or by the lighting scale as a multiple of the shape scale (*relative* lighting scale), we ran a mixed effects logistic regression. The dependent variable was proportion correct, and the independent variables were absolute lighting scale (σ_L), relative lighting scale (σ_L/σ_S), and shape scale (σ_S). Individual differences between observers were modeled as random effects. The regression showed a significant effect of absolute lighting scale ($p < 0.01$), and no significant effect of relative lighting scale ($p = 0.90$) or shape scale ($p = 0.95$). There was a significant interaction between absolute lighting scale and relative lighting scale ($p < 0.001$), and no other significant interactions ($p > 0.05$). Thus performance depends both on absolute lighting scale, and on relative lighting scale by way of an interaction with absolute lighting scale. Both these effects have plausible explanations. Absolute lighting scale affects the spatial frequency content of the image, which as is well known has strong effects on performance in a wide range of tasks (Watson,

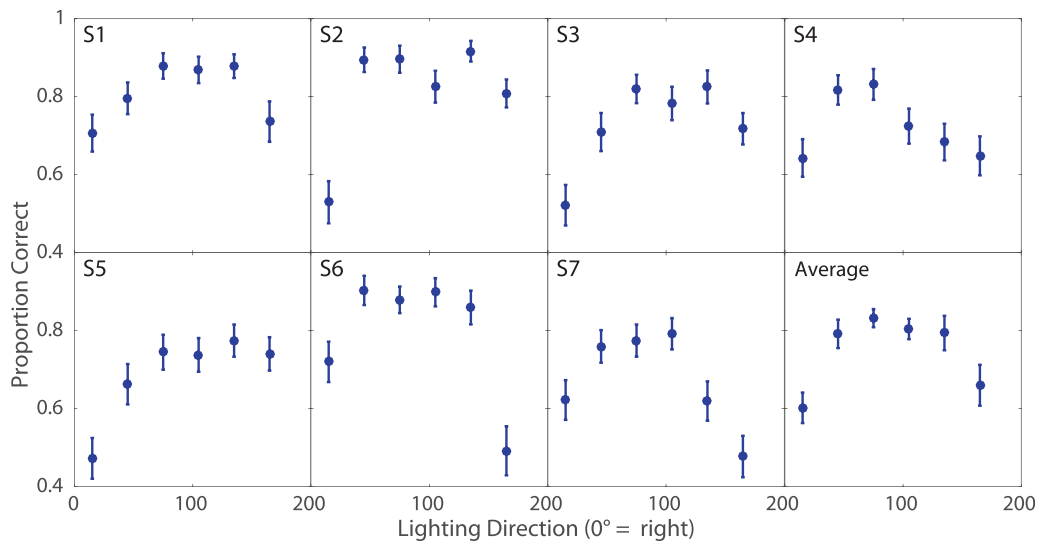


Figure 5. Results from Experiment 1, single lighting direction conditions. Each panel plots proportion correct in the depth discrimination task versus lighting direction. Error bars show 1 SE.

Ahumada, & Farrell, 1986). Relative lighting scale affects the extent to which luminance variations due to shading occur at the same scale as luminance variations due to lighting changes, and so will naturally affect performance on a shape from shading task; we return to this point in the General discussion.

How rapidly can lighting direction vary across a scene before perception of shape from shading breaks down? Figure 4 shows that observers' shape judgments were highly robust to changes in illumination direction. The x coordinates of the three triangles near the bottom of each panel show the values of the three shape scale constants used to generate stimuli in the low, medium, and high frequency shape conditions. When the lighting scale was about the same as the shape scale, observers' performance was still well above chance, and was often not substantially worse than in the single lighting direction condition where $\sigma_L = \infty$. Thus, even when the lighting direction changed on approximately the same spatial scale as the depth of the small bump-like features being judged, observers were still able to make reasonably accurate depth judgments.

Even at low lighting frequencies, our stimuli are intrinsically ambiguous. As for any 2D image, our stimuli are consistent with an infinite set of shape, reflectance, and lighting combinations. At low-lighting frequencies observers perceived the shapes veridically (i.e., in accord with the simulated uniform-reflectance surfaces that were used to generate the stimuli), whereas at high lighting frequencies they perceived a nonveridical combination of shape, reflectance, and lighting (Figure 2, right-hand stimulus), and so their performance on the shape task deteriorated.

These findings show that shape from shading mechanisms can accommodate large and rapid varia-

tions in lighting direction. The lighting variations in our stimuli are not typical of variations in natural scenes, where lighting intensity and diffuseness vary as well as lighting direction, and lighting direction changes are a mix of sudden changes at sharp shadow boundaries and more gradual changes. Rather than emulating natural lighting conditions, our goal was to measure the performance on a shape judgment task under extreme conditions. Many models of shape from shading assume that the visual system infers a single global light source, and our findings challenge such theories; they show that observers are much more flexible than this, and can perceive shape from shading even when lighting direction changes randomly at the scale of the shape features being judged.

The stimuli in Experiment 1 were similar to the shaded disk patterns that have long been used to study the role of the light-from-above prior in perception of shape from shading (Ramachandran, 1988; Morgenstern et al., 2011; Metzger, 1936). Figure 5 shows that the light-from-above prior also affected our observers' percepts. The figure shows performance in the single lighting direction condition, as a function of lighting direction, averaged across the three shape scales, binned into 30° intervals from 0° to 180° . As in studies with shaded disk stimuli, performance was best when light came from approximately overhead, and worst when light came from the left or right. This is expected, as the bumpy surfaces in our stimuli were relatively shallow, so there were few cast shadows and only weak interreflections. As a result, the stimulus images were ambiguous, in that reversing the sign of depth displacements and changing the lighting tilt by 180° would leave the images almost unchanged. The light-from-above prior resolves this ambiguity when the

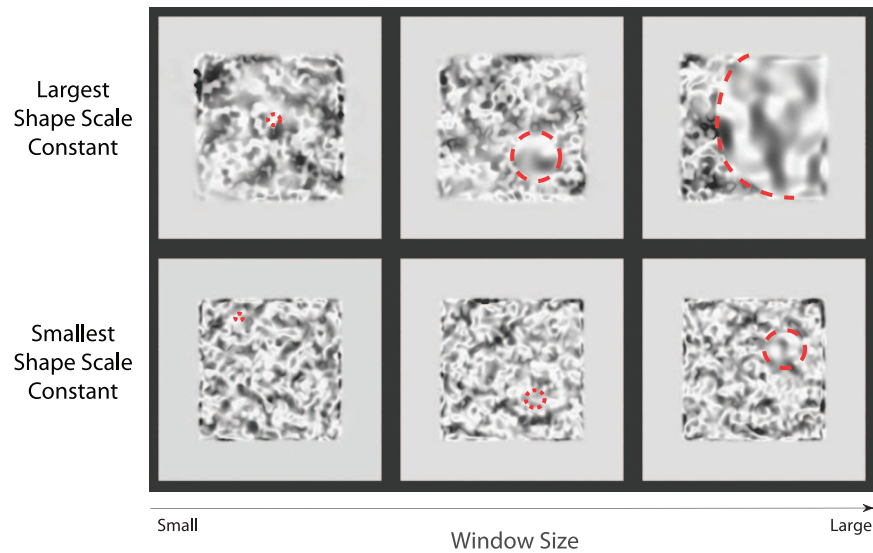


Figure 6. Examples of windowed lighting direction stimuli shown in Experiment 2. The lighting direction changes rapidly from place to place, except within a circular window where the lighting direction is constant. The dashed red lines show the regions of consistent lighting, and were not shown in the experiment. The window size is a multiple of the shape scale constant, so the windows are larger in the top row (large shape scale constant) than in the bottom row (small shape scale constant).

stimulus is consistent with illumination either from overhead or from below. (For reasons that are poorly understood, the peak of the light prior is, on average, slightly to the left of overhead (Sun & Perona, 1998), although this varies substantially across observers (Adams, 2007)). For stimuli rendered with left- or rightward illumination, the lighting prior does not resolve the yoked ambiguities in illumination tilt and curvature sign. As a result, performance drops towards chance for these stimuli.

Thus there was no information, either in the stimulus or in the light-from-above prior, that would allow observers to perform better than chance on trials where the local lighting direction was directly to the left or right. To see whether these ambiguous trials affected our results, we re-ran the logistic regression reported above, after discarding all trials where the local lighting direction at the location of the shape task probes was more than 45° from overhead. The results were the same as before: We found a significant effect of absolute lighting scale (σ_L) on proportion correct ($p < 0.004$), and no significant effect of either relative lighting scale (σ_L/σ_S ; $p = 0.41$) or shape scale (σ_S ; $p = 0.72$). Again there was an interaction between absolute lighting scale and relative lighting scale ($p < 9.2 \times 10^{-8}$), and no other significant interactions ($p = 0.07$).

Experiment 2

In Experiment 1 we found that observers can tolerate rapid variations in local lighting direction when judging

depth based on shading cues. To test the robustness of this finding, in Experiment 2 we measured the minimum size of the region of uniform lighting that observers need in order to perceive shape from shading reliably. Observers made depth judgments on the same type of bumpy surfaces used in Experiment 1. Here though, surfaces were illuminated by highly variable lighting, generated using the highest spatial frequency lighting noise from Experiment 1, except for a circular window around the points whose depth was being judged. Inside this window the lighting direction was constant. We measured performance for several window sizes, to determine the size of the region required for observers to recover shape from shading.

Methods

Observers

There were four observers (two male, two female; mean age 25 years). One was author JW. The others were unaware of the purpose of the experiment and were paid for their participation.

Stimuli

The stimuli were created using the same type of bumpy square panels used in Experiment 1, generated with the same shape scale constants ($\sigma_S = 0.35, 0.53, 0.70$ cm), but new samples of shape noise. Each stimulus was a composite of a single lighting direction image $I_1(x)$, where the lighting direction was the same across the whole image, and a lighting noise image

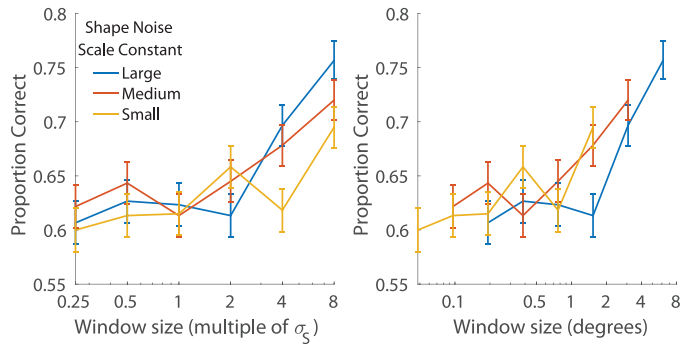


Figure 7. Results from Experiment 2. The figure shows proportion correct in the depth judgment task versus the diameter of the uniform illumination window. The diameter of the window is measured as a multiple of the shape noise constant in each condition (left) or in degrees (right). Error bars show 1 SE.

$I_2(\mathbf{x})$, where the local lighting direction varied randomly across the image as in Experiment 1 (Figure 6). The lighting direction in $I_1(\mathbf{x})$ was chosen randomly and uniformly between 0° and 180° . The lighting noise image $I_2(\mathbf{x})$ was generated using a lighting noise scale constant of $\sigma_L = 0.175$ cm, which was the smallest value used in Experiment 1, and so generated the highest spatial frequency lighting noise. We chose two probe points for the depth judgment, using the same criteria as in Experiment 1. The composite stimulus $I(\mathbf{x})$ was a weighted average of $I_1(\mathbf{x})$ and $I_2(\mathbf{x})$, with practically all weight assigned to $I_1(\mathbf{x})$ inside a circular window centered at the probe points, and practically all weight assigned to $I_2(\mathbf{x})$ outside the circular window. If we denote the midpoint between the two probe points as \mathbf{x}_0 , then the composite stimulus $I(\mathbf{x})$ was given by

$$I(\mathbf{x}) = [1 - w(\mathbf{x})] I_1(\mathbf{x}) + w(\mathbf{x}) I_2(\mathbf{x}),$$

with $w(\mathbf{x}) = \Phi(|\mathbf{x} - \mathbf{x}_0|, d/2, \sigma_w)$

Here $\Phi(x, \mu, \sigma)$ is the normal cumulative distribution function, $|\mathbf{x} - \mathbf{x}_0|$ is the distance between image points \mathbf{x} and \mathbf{x}_0 , d controls the diameter of the circular window, and σ_w (which we set to 0.12 cm) controls the width of the transition region between the two images being combined. The window diameter was set on each trial to 0, 0.25, 0.5, 1, 2, 4, or 8 times the shape scale constant σ_S . The stimuli were shown on the same monitor and at the same viewing distance (57 cm) as in Experiment 1.

Procedure

Each observer completed 12 sessions, where each session included 350 trials and lasted approximately 15 minutes, including a 2–3 min break. There were 21 stimulus conditions, randomly interleaved: seven window sizes crossed with three shape scale constants (see

Stimuli section). The sequence of events on each trial was the same as in Experiment 1, and observers judged which of two red probe points was closer. We generated 40 unique images for each of the 21 stimulus conditions, and each unique image was shown on five separate trials, with the probes at a different location each time. Thus there were $40 \times 5 = 200$ trials per condition, and $21 \times 200 = 4,200$ trials total. Trials were divided randomly among the 12 sessions.

Results and discussion

Figure 7 shows proportion correct, averaged across observers, as a function of the diameter of the uniform lighting window. Performance was better with larger windows. To examine the effect of window size and shape scale on proportion correct, we ran a mixed effects logistic regression. The dependent variable was proportion correct, and the independent variables were the window size in degrees, the window size as a multiple of shape scale, and shape scale. Individual differences between observers were modeled as random effects. Using AIC to penalize model complexity, we found that an interaction term between shape scale and the window size was not justified, so we only tested for an interaction between the two methods of measuring window size. The regression showed a significant effect of window size relative to shape scale ($p < 0.05$). There was no significant effect of window size in degrees ($p = 0.09$), or of shape scale ($p = 0.58$), and there was no interaction between the two methods of measuring window size ($p = 0.52$).

At the smallest window size, performance was about the same ($\sim 60\%$) as in the highest spatial frequency lighting noise condition in Experiment 1, and at the largest window size, performance was about the same ($\sim 72\%$) as in the single lighting direction conditions in Experiment 1. Observers needed a window diameter roughly 2–4 times the shape scale in order for performance to begin to rise above its lowest level ($\sim 60\%$). In Experiment 1, observers required a lighting scale constant that was approximately equal to the shape scale in order for performance to begin to rise above its lowest level. In that experiment, the lighting scale was not the diameter of a window, but rather the scale constant of the Gaussian convolution kernel used to generate the lighting noise. A Gaussian kernel does not have a sharply defined width, but can be roughly described as being 2–4 scale constants wide. Thus the findings of this experiment are broadly consistent with those of Experiment 1, given the differences in how we constructed lighting conditions in the two cases.

These results are consistent with Erens, Kappers, and Koenderink (1993b), who found that observers are unable to judge whether highly local patches of

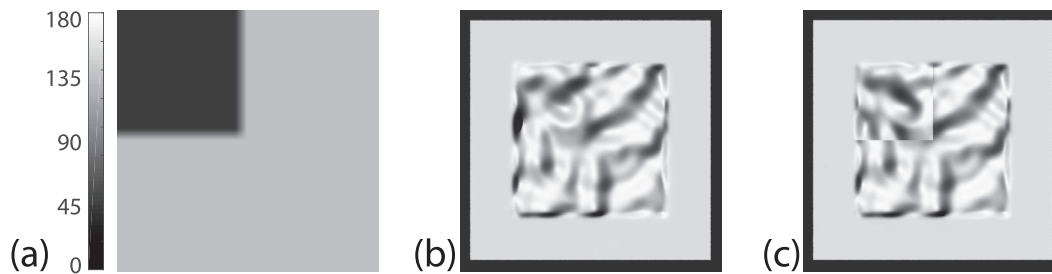


Figure 8. Typical lighting map and stimulus from Experiment 3. (a) A lighting map that indicates a lighting direction of 45° in the top left quadrant and 135° in the other three quadrants, with a rapid but smooth transition between the two regions. (b) A composite stimulus created using the lighting map in panel (a). (c) A composite stimulus created using a lighting map like the one in panel (a), except that here the lighting map is unsmoothed and has a sudden transition between the two regions. The unsmoothed stimulus type shown in panel (c) was not used in the experiment, and is shown here for comparison.

quadratic surfaces are elliptic or hyperbolic. In their study, as in ours, observers perform poorly at a local shape from shading task, and require a window larger than a local quadratic patch in order to perceive shape from shading accurately.

The results of Experiments 1 and 2 are broadly consistent with the findings of van Doorn et al. (2011, 2012). They found that observers perceived simulated bump patterns in accordance with the simulated illumination field when the lighting conditions were typical of real-world scenes (i.e., uniform or divergent), but not when lighting direction varied from place to place either too rapidly (their “random” condition) or in a physically implausible way (e.g., their “circular” conditions). In our experiments too, observers’ shape perception was consistent with the simulated illumination when this was invariant over large enough regions, but not when it varied too rapidly from place to place. However, it is less clear whether our results were consistent with van Doorn et al.’s in terms of what constitutes “too rapidly.” In our experiments, observers were able to judge shape even when the lighting direction varied at approximately the scale of the shape patterns being judged, which is similar to van Doorn et al.’s “random” condition in which simulated bumps were perceived as a mixture of bumps and dents. However, several differences between the two experiments make direct comparisons difficult. For example, in our experiments lighting directions were always above the horizon, whereas in van Doorn et al.’s they spanned 360° ; our observers judged small depth differences between two probe points, and theirs judged whether every one of six shaded disks were seen as bump-like; and our stimuli were smooth Gaussian surfaces, and theirs were six clearly separated bumps, rendered using simple linear gradients. Some combination of these differences may explain apparent discrepancies between the findings.

Experiment 3

In the stimuli used in Experiment 1, it is often difficult to detect changes in lighting direction, except in the highest-frequency lighting conditions where the images cease to look like shaded 3D surfaces (Figure 2). This suggests that observers were able to accommodate rapid changes in lighting direction, even when they are unaware of these changes. In Experiment 3 we pursued this observation. We tested whether observers can detect sudden changes in lighting direction, and whether such changes interfere with performance in a shape from shading task. We divided shape stimuli like those in Experiments 1 and 2 into four quadrants. Three quadrants were illuminated from one direction, and the fourth was illuminated from a different direction (Figure 8). We tested whether observers could identify which quadrant was illuminated differently, and also whether their estimates of shape from shading were less accurate in the odd-one-out quadrant than in the three quadrants lit from the same direction.

Methods

Observers

There were six observers in the lighting task, and four different observers in the shape task. In the shape task, one observer was author JW. All other observers were unaware of the purpose of the study, had not previously seen the stimuli, and were paid for their participation.

Stimuli

The stimuli were created using the same type of bumpy square panels used in Experiments 1 and 2, generated with new samples of shape noise at the medium shape scale constant ($\sigma_S = 0.53$ cm). The

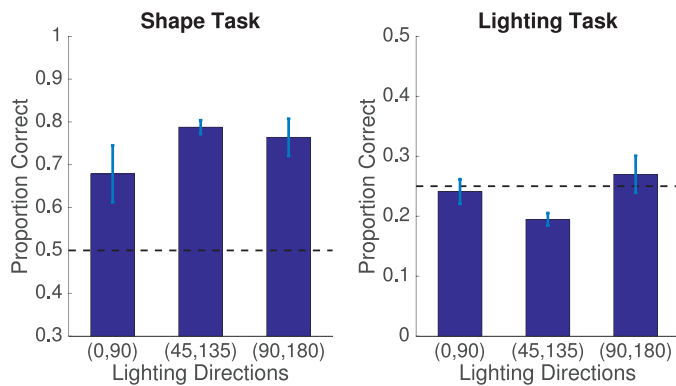


Figure 9. Performance in the shape task (left) and the lighting task (right) of Experiment 3 as a function of the two lighting directions. Chance performance is denoted by the dashed black line. Error bars denote 1 SE.

stimuli were displayed on the same monitor as in Experiments 1 and 2, and at the same viewing distance (57 cm).

As in Experiments 1 and 2, the stimuli were composite images created by combining single lighting direction images pixel by pixel according to a lighting direction map (Figure 8). Here the lighting map indicated one lighting direction in three randomly chosen quadrants, and another lighting direction offset by 90° in the other quadrant, with a rapid but smooth transition between the two regions. We call the three quadrants illuminated from the same direction the *majority* quadrants, and the fourth the *minority* quadrant. The majority or minority region was randomly assigned a lighting direction tilt of 0°, 45°, or 90°, and the other region was assigned a lighting tilt 90° greater. The lighting direction transitioned linearly between the two regions over a boundary strip that was 0.77 cm wide (see Figure 8a). As in Experiments 1 and 2, all lighting directions were 30° off the line of sight, so although the two lighting directions differed by 90° in tilt, they actually formed an angle of 41° relative to each other.

Procedure

In the lighting task, each observer completed one 5-min session of 120 trials. On each trial the stimulus appeared for 1.5 s, followed by a uniform gray screen at the background luminance of the stimulus. We instructed observers to judge which of the four quadrants was illuminated differently from the other three, i.e., which was the minority quadrant. The observer pressed one of four keys to indicate their response. A high or low frequency beep indicated whether the observer's response was correct, and then the next trial began.

The shape task was largely the same as in Experiment 1. Each observer completed one 5-min session of 120 trials. On each trial the observer saw a square panel with two red probe points approximately at the center of one of the four quadrants. On half the trials the probe was in a randomly chosen majority quadrant, and on the other half it was in the minority quadrant. The position of the minority quadrant was chosen randomly on each trial. The positions of the probe points were chosen using the same criteria as in Experiments 1 and 2. The observer pressed one of two buttons to indicate which probe location was closer, the left or the right. The stimulus remained onscreen until the observer responded, and no feedback was given. After the observer responded, there was a short pause and then the next trial began.

Results and discussion

In the shape task, mean proportion correct across observers was not significantly different in the majority quadrants (76.5% correct) and minority quadrant (73.0%; $t = -1.30$, $p = 0.24$). Furthermore, performance was similar to that in the single lighting direction conditions of Experiment 1 (Figure 4). Figure 9 (left panel) shows that performance was substantially above chance for all three lighting direction pairs.

In the lighting task, mean proportion correct across observers (23.5%) was not significantly different from the chance level of 25% correct ($t = -1.12$, $p = 0.31$). Observers performed much worse in the lighting task than in the shape task, even though the stimulus duration in the lighting task (1.5 s) was longer than the mean (1.32 s) and median (1.13 s) response times in the shape task (in which the stimulus was shown until observers responded), and we gave trial-by-trial feedback only in the lighting task. Figure 9 (right panel) shows that performance did not exceed chance for any of the three lighting direction pairs.

These findings support our informal observation that although observers can judge shape from shading under rapid and physically implausible changes in lighting direction, they are often unable to detect those lighting changes. What is striking in the present experiment is that it shows how very rapid lighting changes can be, and still be undetectable. Figure 8b shows a typical stimulus where the lighting direction shifts rapidly but smoothly between the majority and minority regions. Figure 8c shows the same stimulus with a sudden, unsmoothed transition in lighting direction; in this image there are sharp luminance edges at the boundary, and it is clear that the two different lighting directions create very different luminance gradients in the image. However, smoothing this transition over a very narrow region (0.77 cm) is

enough to make a 41° change in lighting direction undetectable. This finding supports and extends the results of Ostrovsky et al. (2005), who found that observers are insensitive to illumination variations within scenes.

Experiment 4

Is it possible that observers were unable to detect the illumination changes in Experiment 3 because the changes were too small, given observers' potentially imprecise estimates of illumination direction? In Experiment 4 we modified the illumination task so that observers were required to detect an illumination change over time, rather than over space, in order to test whether under different stimulus conditions observers could discriminate between the lighting directions we used in Experiment 3.

Methods

Observers

There were four observers. One was author JW, and one other observer was also aware of the purpose of the study. The other two observers were unaware of the purpose of the study, had not previously seen the stimuli, and were paid for their participation.

Stimuli

The stimuli were the single lighting direction stimuli used in Experiment 1, with the medium shape scale constant ($\sigma_S = 0.53$ cm).

Procedure

Each observer completed one 5-min session of 100 trials. On each trial the observer made a 2IFC lighting direction judgment. A surface was shown for 500 ms with one lighting direction, followed by a blank screen for 500 ms, followed by the same surface with a different lighting direction for 500 ms, followed by a blank screen until the observer responded. On half the trials the surface had a lighting direction of 45° in the first interval and 135° in the second interval. On the other half the order was reversed. The observer pressed one of two keys to indicate whether the lighting direction shifted to the left or to the right. A high or low frequency beep indicated whether the observer's response was correct, and then the next trial began.

Results and discussion

Performance levels for the four participants were 96% (author JW), 91%, 91%, and 66% correct, which are all significantly better than chance (50%; $p < 0.05$). On the one hand, the stimuli in this experiment were larger than those in Experiment 3, which might lead us to expect the present task to be easier. On the other hand, the two lighting directions were shown successively instead of simultaneously, which might make the present task harder. In fact observers performed much better in this experiment than in Experiment 3, demonstrating that under some conditions observers can distinguish between lighting directions that differ by 41° in our stimuli. This suggests that observers' inability to detect the illumination inconsistencies in Experiment 3 stems from a specific inability to detect illumination changes over a single surface.

General discussion

In these experiments we investigated perception of shape from shading under lighting that varies from place to place across a surface. In Experiments 1 and 2 we found that observers can make reliable shape judgments as long as the lighting direction is approximately constant over regions on the same scale as the bump-like features being judged. Interestingly, however, Experiment 3 showed that despite robust shape from shading performance, observers were unaware of the substantial changes in illumination within the images. In other words, observers' shape from shading judgments is highly robust to inconsistencies in local lighting direction, but these inconsistencies are often not reportable.

These results are consistent with findings in the related field of lightness perception that show that observers are able to discount local lighting conditions when estimating surface reflectance (e.g., Gilchrist, 2006; Gilchrist & Radonjić, 2010). Different parts of a scene are often illuminated with different intensities, and to disentangle surface reflectance from image luminance the visual system must compensate for variations in lighting intensity. Gilchrist and Radonjić (2010) suggest that the visual system infers lighting boundaries at blurry luminance edges and depth boundaries. The present findings show that observers are similarly able to take highly local lighting conditions into account when judging shape from shading. Furthermore, we find that when estimating shape from shading, the visual system does not need to divide a scene into discrete lighting regions, but instead can accommodate rapidly and smoothly varying lighting conditions. Moreover, our observers were unaware of

the illumination variations within the image. This suggests that observers may implicitly estimate illumination to recover shape and reflectance, without explicit knowledge of the illumination conditions. In line with this hypothesis, Kerrigan and Adams (2013) used visual-haptic training to teach observers implicitly that green and red illumination came from different directions, and observers' subsequent perception of shape from shading was contingent upon illumination color, although observers were completely unaware of the acquired illumination-color relationship.

The fact that the visual system can estimate shape from shading under highly variable lighting conditions, along with the observation that the light-from-above prior plays a role in these estimates (Figure 5), leads naturally to the idea that the visual system may not estimate local lighting conditions at all when estimating shape from shading, and instead may simply rely on the light-from-above prior. This would explain why the visual system seems to have such weak constraints, or even no constraints, on how lighting conditions can vary from place to place. It would also be consistent with recent work on shape from shading, texture, and mirrored surfaces, that suggests that image isophotes (iso-luminance image contours) play an important role in 3D shape perception (Fleming et al., 2011; Kunsberg & Zucker, 2012). Isophotes in some image regions are stable when lighting direction is roughly constant at the scale of the image features being judged, but would presumably become unreliable as a shape cue under very high-frequency lighting changes. These theories, as well as the data presented here, are also consistent with the findings of Erens, Kappers, and Koenderink (1993a), who show that, under some conditions, global shading patterns that provide information about lighting conditions have no influence on observers' performance in a local shape from shading task. This suggests that people sometimes disregard global lighting or shading information when determining local surface shape.

However, several other studies rule out the idea that lighting direction estimates play no role at all in shape from shading. Ramachandran (1988) found that when observers view an array of ambiguous shaded disks that are consistent with lighting from the left or right, the perceived lighting direction switches between left and right for all disks at once; observers do not simultaneously see some disks illuminated from the left and others from the right. Similarly, Morgenstern, Murray, and Harris (2011) asked observers to judge the shape of ambiguous shaded disks embedded in scenes where there were strong lighting direction cues. Observers reported the disks as having whichever shape was consistent with the lighting direction cues, even when these cues were several degrees of visual angle away and on a different object. Both these studies show that, in

the absence of contradictory image information, the visual system tends to interpret scenes in accordance with a single lighting direction. Furthermore, Adams, Graf, and Ernst (2004) showed that when observers make shape from shading judgments for 90 min in an environment where light comes from some direction other than directly above, their prior over lighting directions shifts towards the illumination conditions experienced in the training environment. This trained shift even affects performance on a separate lightness judgment task. Of course isophotes may still play an important role in 3D shape perception, but these studies show that remaining ambiguities may be resolved by imposing spatial consistency across lighting direction estimates.

Koenderink and Pont (2003) proposed a biologically plausible method for estimating the tilt component of lighting direction (i.e., the projection of the lighting direction into the image plane) using the first and second partial derivatives of the luminance image. To test whether the falloff in performance that we observed with high frequency lighting noise (Figure 4) is consistent with such an estimator, we implemented a local version of their method. Instead of averaging information from derivatives (specifically, their G_2 matrix) over the entire image, we used a local weighted average with weights given by a Gaussian window with scale constant $\sigma_{KP} = 1.2^\circ$, centered at the image location where the lighting direction is being evaluated. Other choices for the size of the Gaussian window, between $\sigma_{KP} = 0.12^\circ$ and $\sigma_{KP} = 8^\circ$ give similar results. We used this method to estimate the lighting direction in the same stimuli we showed to human observers, at the same test locations where observers judged relative depth. Figure 10 shows that lighting direction estimates were most accurate with low frequency lighting noise (large scale constant σ_L), and became less reliable when the lighting noise scale was smaller than the estimator's integration window. Thus our behavioral findings are qualitatively consistent with the idea that observers perform worse under lighting noise because such noise makes it more difficult to estimate the local lighting direction. To go further and model proportion correct in the relative depth judgment task would require integrating Koenderink and Pont's estimator into a more complete model of shape from shading, which we leave for future work.

How do we resolve the seeming inconsistency between the present experiments, where observers did not seem to infer a consistent lighting direction across a scene, and Ramachandran (1988) and Morgenstern et al. (2011), where they did? In Ramachandran's and Morgenstern et al.'s experiments, the regions between ambiguous shaded patterns were empty, or at least contained no cues to lighting direction: Shaded disks were separated by empty or ambiguous regions. In

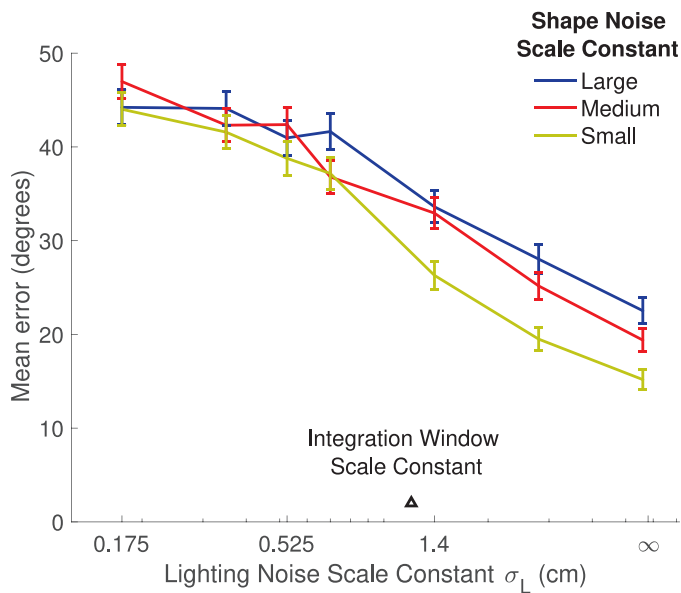


Figure 10. The mean absolute error of the estimates of the light source direction, using the method described in Koenderink and Pont (2003), as a function of lighting noise scale constant. The black triangle shows the scale constant of the Gaussian integration window of the lighting direction estimator.

contrast, stimuli in the present experiments were composed of bumpy surfaces, and each bump-like feature was consistent with just two opposed lighting directions (assuming constant surface reflectance). In the conditions where lighting direction varied from place to place, distant image locations were separated by bump-like features that provided local lighting direction cues. In these conditions, it would make little sense for the visual system to infer a single lighting direction across large stimulus regions, as this would represent an implausible interpretation of the image. Instead, we suggest that the human visual system favors image interpretations that are consistent with a single global lighting direction, but combines this preference, or prior, with other sources of information. Consistent with this view, Adams and Elder (2014) showed that, when viewing multiple shaded objects, observers' shape judgments were consistent with a prior for global illumination, but that this prior competes with other priors (e.g., for an overhead lighting direction, and for object convexity) as well as image cues (e.g., the presence of specular highlights, which promote a percept of convexity).

This view accommodates the fact that human vision assumes that illumination is most likely to come from overhead, and also that this prior can be overruled by lighting cues elsewhere in the scene (Morgenstern et al., 2011) or by information from disparity (Adams, Kerrigan, & Graf, 2010) or touch (Adams et al., 2004) that disambiguates both shape and illumination direction. Thus the visual system uses local lighting direction

cues to estimate shape from shading, but can also recruit additional information (from priors, from neighboring visual regions, or from other modalities) when local lighting direction cues are unavailable or ambiguous. Information from other sources (e.g., from priors or distant lighting cues) is likely to be combined with local lighting information in a statistically rational way, as shown by Morgenstern et al. (2011) and Adams and Elder (2014).

Our motivation for this study was to investigate how the human visual system uses lighting cues to infer three-dimensional shape. We find that shape from shading mechanisms are highly flexible, accommodating substantial variations in the direction of infinitely distant point-light sources. Strikingly, perception of shape from shading only breaks down when the lighting direction changes at scales that are equal to or smaller than the shape features being judged. We have not carried out an ideal observer analysis of these experiments (Watson, 1987; Geisler, 1989), but it seems likely that an ideal observer would show a similar pattern of performance, since lighting variations at the same scale as depth modulations disrupt the signal used by shape from shading mechanisms, namely luminance variations at the scale of shape variations. If this view is correct, then even the breakdown in performance that we observed at high lighting noise frequencies is at least partly due to a reduction in available information, rather than solely reflecting limitations in the perceptual representation of lighting. In any case, our findings show that the visual system's use of lighting cues is highly flexible, to the extent that it can accommodate a high degree of lighting variability that may exceed the variations we encounter in most real scenes.

Keywords: shape from shading, shape from X, illumination, depth discrimination

Acknowledgments

JDW was funded by an NSERC CREATE grant. RFM was funded by an NSERC Discovery Grant.

Commercial relationships: none.
Corresponding author: John D. Wilder.
Email: jdwilder@cs.toronto.edu.
Address: Department of Psychology, University of Toronto, Toronto, Canada.

References

Adams, W. J. (2007). A common light-prior for visual search, shape, and reflectance judgments. *Journal of*

- Vision*, 7(11):11, 1–7, <https://doi.org/10.1167/7.11.11>. [PubMed] [Article]
- Adams, W. J., & Elder, J. H. (2014). Effects of specular highlights on perceived surface convexity. *PLoS Computational Biology*, 10(5), 1–13.
- Adams, W. J., Graf, E. W., & Ernst, M. O. (2004). Experience can change the ‘light-from-above’ prior. *Nature Neuroscience*, 7(10), 1057–1058.
- Adams, W. J., Kerrigan, I. S., & Graf, E. W. (2010). Efficient visual recalibration from either visual or haptic feedback: The importance of being wrong. *Journal of Neuroscience*, 30(44), 14745–14749.
- Barron, J. T., & Malik, J. (2015). Shape, illumination, and reflectance from shading. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37(8), 1670–1687.
- Basri, R., & Jacobs, D. W. (2003). Lambertian reflectance and linear subspaces. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 25(2), 218–233.
- Belhumeur, P. N., Kriegman, D. J., & Yuille, A. L. (1999). The bas-relief ambiguity. *International Journal of Computer Vision*, 35(1), 33–44.
- Bloj, M., Ripamonti, C., Mitha, K., Hauck, R., Greenwald, S., & Brainard, D. H. (2004). An equivalent illuminant model for the effect of surface slant on perceived lightness. *Journal of Vision*, 4(9):6, 735–746, <https://doi.org/10.1167/4.9.6>. [PubMed] [Article]
- Boyaci, H., Maloney, L. T., & Hersh, S. (2003). The effect of perceived surface orientation on perceived surface albedo in binocularly viewed scenes. *Journal of Vision*, 3(8):2, 541–553, <https://doi.org/10.1167/3.8.2>. [PubMed] [Article]
- Cuttle, C. (2008). *Lighting by design*. New York, NY: Routledge.
- Dror, R. O., Willsky, A. S., & Adelson, E. H. (2004). Statistical characterization of real-world illumination. *Journal of Vision*, 4(9):11, 821–837, <https://doi.org/10.1167/4.9.11>. [PubMed] [Article]
- Erens, R. G., Kappers, A. M., & Koenderink, J. J. (1993a). Estimating local shape from shading in the presence of global shading. *Perception & Psychophysics*, 54(3), 334–342.
- Erens, R. G., Kappers, A. M., & Koenderink, J. J. (1993b). Perception of local shape from shading. *Perception & Psychophysics*, 54(2), 145–156.
- Fleming, R. W., Holtmann-Rice, D., & Bülthoff, H. H. (2011). Estimation of 3d shape from image orientations. *Proceedings of the National Academy of Sciences, USA*, 108(51), 20438–20443.
- Geisler, W. S. (1989). Sequential ideal-observer analysis of visual discriminations. *Psychological Review*, 96(2), 267–314.
- Gershun, A. (1939). The light field. *Studies in Applied Mathematics*, 18(1–4), 51–151.
- Gilchrist, A. (2006). *Seeing black and white*. Oxford, UK: Oxford University Press.
- Gilchrist, A., Kossyfidis, C., Bonato, F., Agostini, T., Cataliotti, J., Li, X., . . . Economou, E. (1999). An anchoring theory of lightness perception. *Psychological Review*, 106(4), 795–834.
- Gilchrist, A. L., & Radonjić, A. (2010). Functional frameworks of illumination revealed by probe disk technique. *Journal of Vision*, 10(5):6, 1–12, <https://doi.org/10.1167/10.5.6>. [PubMed] [Article]
- Horn, B. (1975). Obtaining shape from shading information. In P. H. Winston (Ed.), *The psychology of computer vision* (115–155). New York, NY: McGraw-Hill.
- Kartashova, T., Sekulovski, D., de Ridder, H., te Pas, S. F., & Pont, S. C. (2016). The global structure of the visual light field and its relation to the physical light field. *Journal of Vision*, 16(10):9, 1–18, <https://doi.org/10.1167/16.10.9>. [PubMed] [Article]
- Kerrigan, I. S., & Adams, W. J. (2013). Learning different light prior distributions for different contexts. *Cognition*, 127(1), 99–104.
- Koenderink, J. J., & Pont, S. C. (2003). Irradiation direction from texture. *Journal of the Optical Society of America (JOSA) A*, 20(10), 1875–1882.
- Koenderink, J. J., Pont, S. C., van Doorn, A. J., Kappers, A. M. L., & Todd, J. T. (2007). The visual light field. *Perception*, 36(11), 1595–1610.
- Kunsberg, B., & Zucker, S. W. (2012). The differential geometry of shape from shading: Biology reveals curvature structure. In *Computer vision and pattern recognition workshops 2012* (pp. 39–46). Washington, DC: IEEE.
- Langer, M. S., & Bülthoff, H. H. (2000). Depth discrimination from shading under diffuse lighting. *Perception*, 29(6), 649–660.
- Metzger, W. (1936). *Laws of seeing*. Cambridge, MA: MIT Press.
- Moon, P., & Spencer, D. E. (1981). *The photic field*. Cambridge, MA: MIT Press.
- Morgenstern, Y., Geisler, W. S., & Murray, R. F. (2014). Human vision is attuned to the diffuseness of natural light. *Journal of Vision*, 14(9):15, 1–18, <https://doi.org/10.1167/14.9.15>. [PubMed] [Article]
- Morgenstern, Y., Murray, R. F., & Harris, L. R. (2011). The human visual system’s assumption that light comes from above is weak. *Proceedings of the*

- National Academy of Sciences, USA*, 108(30), 12551–12553.
- Mury, A. A., Pont, S. C., & Koenderink, J. J. (2007). Light field constancy within natural scenes. *Applied Optics*, 46(29), 7308–7316.
- Mury, A. A., Pont, S. C., & Koenderink, J. J. (2009a). Representing the light field in finite three-dimensional spaces from sparse discrete samples. *Applied Optics*, 48(3), 450–457.
- Mury, A. A., Pont, S. C., & Koenderink, J. J. (2009b). Structure of light fields in natural scenes. *Applied Optics*, 48(28), 5386–5395.
- Ostrovsky, Y., Cavanagh, P., & Sinha, P. (2005). Perceiving illumination inconsistencies in scenes. *Perception*, 34(11), 1301–1314.
- Ramachandran, V. S. (1988). Perception of shape from shading. *Nature*, 331, 163–166.
- Ramamoorthi, R., & Hanrahan, P. (2001). On the relationship between radiance and irradiance: Determining the illumination from images of a convex Lambertian object. *Journal of the Optical Society of America A*, 18(10), 2448–2459.
- Sun, J., & Perona, P. (1998). Where is the sun? *Nature Neuroscience*, 1(3), 183–184.
- te Pas, S. F., Pont, S. C., Dalmaijer, E. S., & Hooge, I. T. C. (2017). Perception of object illumination depends on highlights and shadows, not shading. *Journal of Vision*, 17(8):2, 1–15, <https://doi.org/10.1167/17.8.2>. [PubMed] [Article]
- Tian, Y.-l., Tsui, H., Yeung, S., & Ma, S. (1999). Shape from shading for multiple light sources. *Journal of the Optical Society of America A*, 16(1), 36–52.
- Todd, J. T., & Mingolla, E. (1983). Perception of surface curvature and direction of illumination from patterns of shading. *Journal of Experimental Psychology: Human Perception and Performance*, 9(4), 583–595.
- van Doorn, A. J., Koenderink, J. J., Todd, J. T., & Wagemans, J. (2012). Awareness of the light field: the case of deformation. *i-Perception*, 3(7), 467–480.
- van Doorn, A. J., Koenderink, J. J., & Wagemans, J. (2011). Light fields and shape from shading. *Journal of Vision*, 11(3):21, 1–21, <https://doi.org/10.1167/11.3.21>. [PubMed] [Article]
- Ward Larson, G., & Shakespeare, R. A. (2004). *Rendering with radiance: The art and science of lighting visualization*. Charleston, SC: BookSurge, LLC.
- Watson, A. B. (1987). The ideal observer concept as a modeling tool. In *Frontiers of visual science: Proceedings of the 1985 symposium* (pp. 32–37). Washington, DC: National Academy Press.
- Watson, A. B., Ahumada, A. J., & Farrell, J. E. (1986). Window of visibility: A psychophysical theory of fidelity in time-sampled visual motion displays. *JOSA A*, 3(3), 300–307.
- Xia, L., Pont, S. C., & Heynderickx, I. (2016). Effects of scene content and layout on the perceived light direction in 3d spaces. *Journal of Vision*, 16(10):14, 1–13, <https://doi.org/10.1167/16.10.14>. [PubMed] [Article]
- Zhang, R., Tsai, P.-S., Cryer, J. E., & Shah, M. (1999). Shape from shading: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 21(8), 690–706.