

# The influence of shape cues on the perception of lighting direction

**James P. O’Shea**

Vision Science Program, University of California,  
Berkeley, CA, USA



**Maneesh Agrawala**

Department of Computer Science, University of California,  
Berkeley, CA, USA



**Martin S. Banks**

Vision Science Program, Department of Psychology and  
Wills Neuroscience Institute, University of California,  
Berkeley, CA, USA



Three scene properties determine the luminances in the image of a shaded object: the material reflectance, the illuminant position, and the object’s shape. Because all three properties determine the image, one cannot solve for any one property without knowing the other two. Nevertheless, people perceive consistent 3D shape and consistent lighting in shaded images; they must therefore be making assumptions about the unknown properties. We conducted two psychophysical experiments to determine how viewers use shape information to estimate the lighting direction from shaded images. In the first experiment, we confirmed that observers use 3D shape information when estimating lighting direction. In the second experiment, we investigated how different shape cues affect lighting direction estimates. Observers can accurately determine lighting direction when a host of shape cues specify the objects. When shading is the only cue, observers always set lighting direction to be from above. We modeled the results in a Bayesian framework that included a prior distribution describing the assumed lighting direction. The estimated prior was slightly counterclockwise from above at a  $\sim 30^\circ$  slant. Our model showed that an assumption of convexity provides an accurate estimate of lighting direction when the shape is globally, but not locally, consistent with convexity.

Keywords: lighting, shape from shading, depth perception

Citation: O’Shea, J. P., Agrawala, M., & Banks, M. S. (2010). The influence of shape cues on the perception of lighting direction. *Journal of Vision*, 10(12):21, 1–21, <http://www.journalofvision.org/content/10/12/21>, doi:10.1167/10.12.21.

## Introduction

The variation in luminance across the image of a surface provides information about the 3D shape of the surface, the material of which the surface is composed, and the lighting or illumination falling on the surface. Although the physics of light transport is well understood (Kajiya, 1986), it remains unclear how human observers estimate reflectance, lighting, and shape from a single image. Consider the simple case of Lambertian reflectance, a single distant light source that is always visible and within  $90^\circ$  of any surface normal, and no inter-reflections. Under these assumptions, the local diffuse shading equation is

$$I(x, y) = \rho(\vec{N}(x, y) \cdot \vec{L}), \quad (1)$$

where  $\rho$  is the constant albedo,  $\vec{N}(x, y)$  is the surface normal at point  $(x, y)$  in the image, and  $\vec{L}$  is the vector pointing toward the light source. Thus, the observed luminance in the image is determined by the reflective properties of the material and the orientation of the

surface normal relative to the lighting direction. Because shape, lighting direction, and material properties all determine the observed image, one cannot in general solve for any one of those properties without knowing the other two. The problem of solving for any of these terms (shape, lighting direction, or reflectance) is consequently ill-posed.

Human observers have stable percepts of lighting direction (Koenderink, Pont, van Doorn, Kappers, & Todd, 2007) and stable percepts of shape (Koenderink, van Doorn, Christou, & Lappin, 1996) in shaded images, which makes sense if the object’s shape or the lighting direction is known, respectively. However, observers also have stable percepts of lighting direction when the shape is not known (Koenderink, van Doorn, & Pont, 2004) and stable percepts of shape when the lighting direction is not known (Ho, Landy, & Maloney, 2006). The latter observations suggest that observers are making assumptions about the two unknown properties in order to solve for the third. Here we investigate how observers use the available sensory data and assumptions to turn an ill-posed problem of estimating lighting direction into a solvable one.

## Models

### Shape-based method

The lighting information contained in shaded objects is illustrated in Figure 1. On the left is an irregular object illuminated by a distant point light. Assume that the observer knows the surface shape. Then one can determine how the observed luminance varies as a function of the 3D orientation of different parts of the object. On the right, we plot luminance as a function of surface slant and tilt, where slant  $\varphi$  is the angle between the line of sight and the surface normal, and tilt  $\theta$  is the direction of that angle relative to horizontal (Stevens, 1983). The plot is regular with a clear peak. Assume also that the surface material is Lambertian and the reflectance is constant. Then from Equation 1, the luminance at each point in the luminance map on the right is informative about the direction of the light source. For instance, the point of maximum luminance has a surface normal pointing toward the light. Points with half the maximum luminance have surface normals that point  $60^\circ$  ( $\cos(60^\circ) = 1/2$ ) away from the light. Assuming that the surface albedo is constant and that the surface slant and tilt are estimated with some degree of accuracy, we can rewrite Equation 1 as the product of two vectors:

$$I(x, y) \propto \vec{N}(x, y)^T \vec{L}, \tag{2}$$

where  $\vec{N} = \begin{bmatrix} N_x \\ N_y \\ N_z \end{bmatrix}$ ,  $\vec{L} = \begin{bmatrix} L_x \\ L_y \\ L_z \end{bmatrix}$ ,

and  $\|\vec{N}\| = \|\vec{L}\| = 1$ .

Using the positive luminance values ( $I(x, y) > 0$ ), we can set up a linear system of equations:

$$A\vec{L} = B, \tag{3}$$

where  $A = \begin{bmatrix} N_x(x_1, y_1) & N_y(x_1, y_1) & N_z(x_1, y_1) \\ N_x(x_2, y_2) & N_y(x_2, y_2) & N_z(x_2, y_2) \\ \vdots & \vdots & \vdots \\ N_x(x_k, y_k) & N_y(x_k, y_k) & N_z(x_k, y_k) \end{bmatrix}$

and  $B = \begin{bmatrix} I(x_1, y_1) \\ I(x_2, y_2) \\ \vdots \\ I(x_k, y_k) \end{bmatrix}$  for  $k$  positive points in the image.

We can then estimate the lighting direction  $\vec{L}$  using a linear least-squares approach:

$$L = (A^T A)^{-1} A^T B. \tag{4}$$

Thus, we can solve for the lighting direction when surface shape and reflectance are known. To estimate the lighting direction in this way, the visual system must measure the luminances and the 3D orientations of points on the object (and the assumption of constant Lambertian reflectance must be valid). Because this method depends on knowing or estimating the 3D surface geometry, we refer to it as the shape-based method for estimating lighting direction. If the luminance and orientation measurements are erroneous, the estimate of lighting direction will be correspondingly erroneous.

### Image-based method

Pentland (1982) developed one of the first image-based methods for estimating lighting direction in an image. Precluding any direct estimation of the 3D geometry of the surface, but assuming the surface normals are isotropically distributed, he showed that the slant and tilt of the

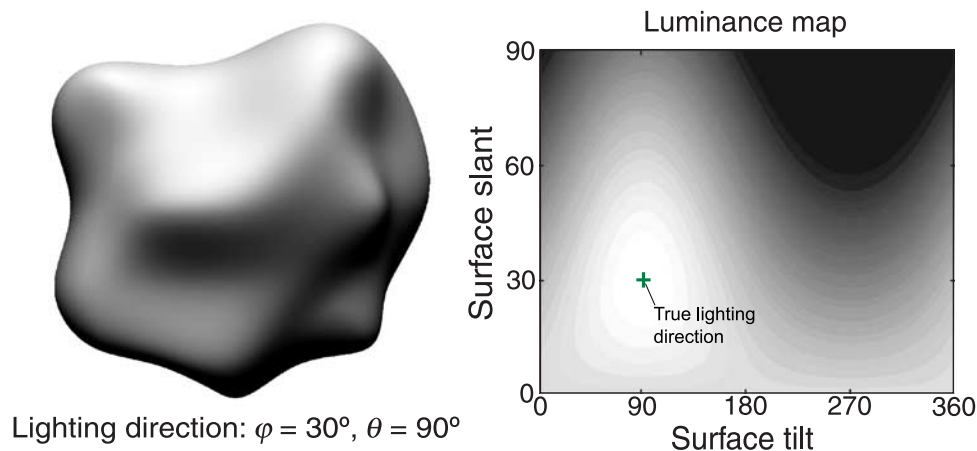


Figure 1. Shaded object and its associated luminance map. The object (left) is illuminated from above (light slant  $\varphi = 30^\circ$ , light tilt  $\theta = 90^\circ$ ). The luminance map plots the observed luminance value for each surface orientation in the image. The green cross indicates the true lighting direction.

lighting direction can be estimated from the local shading derivatives. Lee and Rosenfeld (1985) extended and improved the method, deriving the following estimate of lighting tilt  $\theta$ :

$$\theta = \arctan\left(\frac{E(I_y)}{E(I_x)}\right), \quad (5)$$

where  $E(I_y)$  is the maximum-likelihood estimate of the image derivative along the  $y$  direction and  $E(I_x)$  is the maximum-likelihood estimate of the image derivative along the  $x$  direction. The slant  $\phi$  of the lighting direction is

$$\phi = \arccos\left(\frac{3E(I^2)}{\lambda^2\rho^2} - 1\right), \quad (6)$$

where  $\lambda$  is the illumination brightness,  $\rho$  is the surface albedo, and  $E(I^2)$  is the expectation of  $I^2$  taken along the tilt direction  $\theta$ . The slant and tilt estimates determine the lighting direction as follows:

$$L = \begin{bmatrix} \sin(\theta)\cos(\phi) \\ \sin(\theta)\sin(\phi) \\ \cos(\phi) \end{bmatrix}. \quad (7)$$

Thus, we can solve for the lighting direction using the 2D content of the image and an estimate of albedo and illumination brightness, provided that the global shape of the surface is convex. Because this approach is based on the 2D image information, we refer to it as the image-based method.

## Outline

In this paper, we first investigate whether human observers use a shape-based or image-based approach to estimate the lighting direction in a scene. We then examine humans’ ability to infer lighting direction when the material property was provided and shape information was indicated to greater or lesser extents. In so doing, we learn more about the computations and assumptions that viewers make while interpreting shaded images. When the sensory data (i.e., the image content) reliably specified the 3D shape, the observer could in principle determine the lighting direction (Equation 4). When the sensory data did not specify the 3D shape well, the observer could not determine the lighting direction from those data directly; in that case, the observer had to make assumptions about the shape and the lighting, and those assumptions were presumably based on previous experience.

In our analysis of the results, we use a Bayesian framework to fit a model to the data. We show that a simple model incorporating information from the sensory data and expectations based on previous experience fits

the data well. The best-fitting model implies that observers use the sensory data and prior expectations, but that they rely on the prior expectations when the sensory data are unreliable. By making observers familiar with the material properties and manipulating the shape information, we were able to determine the usefulness of various shape cues and to measure the direction and variance of the prior for lighting direction.

## General methods

### Apparatus

We displayed stimuli on a custom stereoscope with two arms that rotated about vertical axes co-linear with the rotation axes of the eyes (Backus, Banks, van Ee, & Crowell, 1999). Each arm used a mirror to position the image from a CRT in front of each eye. The CRTs were ViewSonic G225f displays with a resolution of  $1280 \times 1024$ . The physical distance from each eye to the appropriate CRT was 39 cm, so each pixel subtended  $2.8 \times 2.6$  arcmin. We gamma-corrected each display to linearize the luminance function for the grayscale images. Except for the CRTs, the room was dark.

The observers stabilized their head position using a bite bar fastened to an adjustable mount. We adjusted the separation of the rotation centers of the stereoscope arms to match each observer’s inter-ocular distance. We rotated the arms so that the vergence angle matched the viewing distance of 39 cm.

### Lighting parameters

We parameterized lighting direction in terms of slant and tilt in much the way Stevens (1983) described surface slant and tilt. Lighting slant ( $\phi$ ) is the angle between a line from the eye (or cyclopean eye) to the center of the object and a vector from the center of the object to the light. Lighting tilt ( $\theta$ ) is the angle between the horizontal axis and the projection of the lighting direction onto the frontal plane. The slant and tilt of the light correspond, respectively, to the zenith and azimuth of the light (Lopez-Moreño, Hadap, Reinhard, & Gutierrez, 2009).

### Stimuli and procedure

On each trial, we presented two stimuli simultaneously: a test object and a response object (Figure 2). The two objects were rendered using the OpenGL graphics library and C++. The objects were composed of matte material (i.e., Lambertian reflectance). We told the observers that the material was similar to matte paper. We also showed

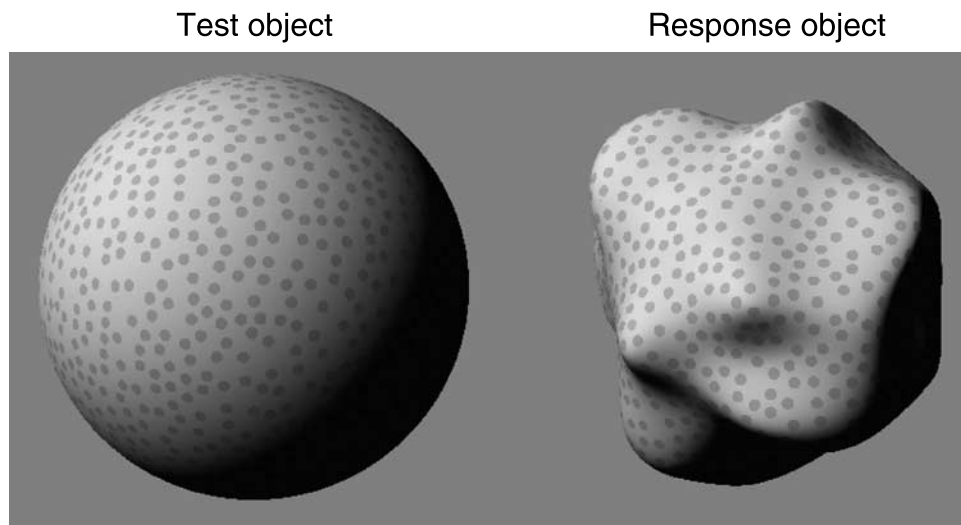


Figure 2. Experimental stimuli. The test object was rendered with different sets of shape cues in different conditions; in the example above, all cues were present including having the familiar shape of a sphere. The response object was always rendered with all shape cues (disparity, texture gradient, global convexity, occluding contour, and shading). Observers adjusted the direction of the light on the response object until it appeared to be the same as the light illuminating the test object.

them that the luminance does not vary as a function of viewpoint. We made clear that the stimuli were composed of the same material as used in the demonstration.

The response object contained all of the shape cues under investigation. We chose to examine a set of shape cues commonly found in the shape perception literature.

1. *Shading*: Each pixel was given the appropriate luminance given the object's shape, Lambertian reflectance, and the direction of the light source at infinite distance. The shading method correctly generated attached shadows. Cast shadows were not rendered, but the lighting slant was never greater than  $45^\circ$ , so there would have been few such shadows anyway.
2. *Texture gradient*: We applied a texture by rendering small gray disks on the object. The disks were oriented in the tangent plane of the surface. We positioned the disks using a dart-throwing algorithm to generate a Poisson-disk sampling of the vertices (Dunbar & Humphreys, 2006). The texture gradient fully specified the 3D shape of the object up to an unknown scale factor. Because the light source is at infinite distance, the scale factor does not need to be known to estimate the lighting direction, so the texture provided the information required to estimate lighting direction accurately (Equation 4).
3. *Binocular disparity*: Each point on the object was given the appropriate horizontal and vertical disparities for the specified shape. This cue fully specified the 3D shape of the object and therefore provided the information needed to estimate light direction accurately (Equation 4).

4. *Global convexity*: When this cue was present, the object was a sphere with random radial perturbations. Because the object was approximately spherical, the orientation at a point on the surface was highly correlated with the point's position in the image. For example, surface points above and to the right of the center of the image had tilts on average of  $45^\circ$ . Thus, observers could in principle have used such regularity to estimate light direction (Lopez-Moreño et al., 2009). The fact that the surface was globally convex is consistent with the convexity assumption observers tend to make about surfaces (Langer & Bülthoff, 2001; Mamassian & Landy, 1998).
5. *Occluding contour*: When this cue was present, the silhouette of the object was visible. The silhouette provides information about 3D shape (Ikeuchi & Horn, 1981; Malik & Maydan, 1989). The slant of the surface at the occluding contour is  $90^\circ$  because that part of the surface is by definition orthogonal to the viewing direction. The tilt is equal to the orientation of the tangent to the contour at that point. Because surface orientation is known at the occluding contour, luminance values along the contour could provide useful information about the lighting direction. Of course, the surface at the occluding contour is invisible to the viewer, so one cannot measure luminance at precisely that point, but one can estimate the luminance by extrapolating from nearby points (Nillius & Eklundh, 2001). From these measurements, one can estimate the tilt of the light: Specifically, the tilt is perpendicular to the orientation of the contour at the brightest point on the

occluding contour. We can see this relationship in the luminance map in [Figure 1](#) where the brightest portion of the map at large surface slants indicates the light tilt.

We generated the 3D objects by subdividing the triangles of an initial control mesh. We created the spheres by subdividing an icosahedron (20-sided regular polyhedron) and normalizing the vertices to be equidistant from the origin of the object. The final spheres were composed of approximately 25,000 triangles. The irregular shapes were created using an implementation of the Catmull–Clark subdivision surfaces algorithm that generates smooth surfaces with  $C^1$  continuity (Catmull & Clark, 1978). The resulting objects were spheres with random radial perturbations ([Figures 2 and 3](#)). We created the surface perturbations by randomly displacing the position of each vertex prior to the third iteration of the subdivision. We continued to run the subdivision algorithm until each object consisted of approximately 100,000 polygons. The globally flat shapes were generated from a planar control mesh. The resulting objects were planes with random perturbations in depth.

We illuminated the test and response objects with point light sources at infinite distance, one source for the test object and another for the response. We told the observers that the light was at infinite distance and thus similar to the sun. Observers moved a trackball to adjust the 2D orientation ( $\varphi, \theta$ ) of the light on the response object. Their task was to make the lighting direction on the response object match the perceived lighting direction on the test object. Lighting direction was not changed online with the trackball movement. Instead, after adjusting the trackball, observers clicked a button to update the lighting on the response object. Thus, they could not see changes in shading due to movement of the light source and, therefore, could not use light motion as an additional cue

to shape. They kept making adjustments until the perceived lighting directions on the response and test objects were the same. They indicated that they were the same by clicking a mouse button. The test object appeared on the left for half the trials and on the right for the other half. A new pair of objects appeared on each trial. The shape of the response object was always well specified, so observers should have accurately perceived its shape regardless of how accurately they perceived the shape of the test object.

To measure the perceived direction of the illuminant, we could conceivably have used an estimation procedure such as asking observers to indicate light direction with a pointer. We chose not to use this approach because we had no way of knowing the mapping between perceived direction and pointer orientation, the so-called *response-mapping problem*. Said another way, one cannot know from the responses of such an estimation procedure which effects are due to the mapping between the percept and the response and which effects are directly indicative of the percept. By focusing on perceptual equivalence, we can be more confident that our results reflect perceptual processes.

## Experiment 1

We first investigated whether observers use a shape-based or image-based approach to estimate the lighting direction in a scene. To do so, we displayed irregular test objects that were globally concave and varied the disparity information specifying the 3D shape. The response object was always globally convex.

First, consider the predictions for image-based methods. If the lighting direction on the test and response objects were the same, the shading patterns on the two objects

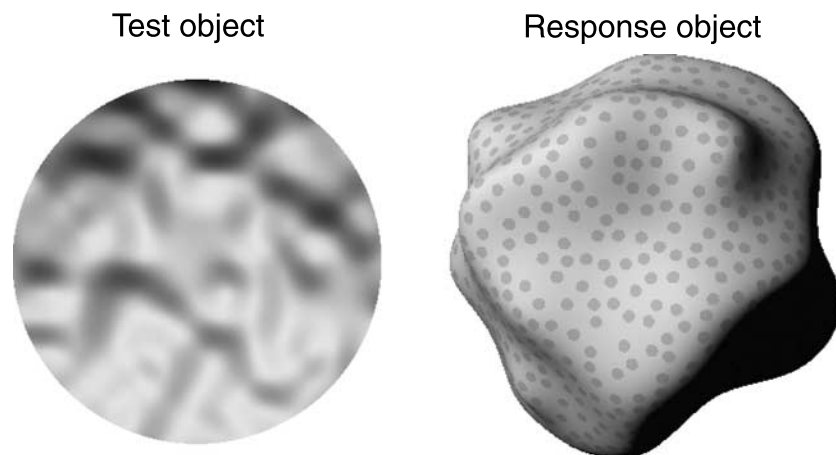


Figure 3. Examples of stimuli used in [Experiment 1](#). The test surface was always globally concave, and we varied the disparity information used to specify its 3D shape. The response object was convex. Observers adjusted the direction of the lighting on the response object until it appeared to be the same as the lighting direction on the test object.

would be in opposite directions (Figure 3). For example, if the lighting on both objects was from above, the response object would be brighter on the top than the bottom and the test object would be brighter on the bottom than the top. To match the shading patterns, the observer would have to set the tilt of the lighting direction on the response object  $180^\circ$  from the lighting tilt on the test object. Thus, image-based methods should yield tilt errors of  $180^\circ$ .

Now consider the predictions for the shape-based method. We manipulated the information specifying the test object's 3D shape by setting the disparities to zero (specifying a flat surface) or to the correct values for the shape (specifying a concave surface). With zero disparities, the shape of the test object was ambiguous, and observers generally perceived the shape as globally convex (Langer & Bühlhoff, 2001; Mamassian & Landy, 1998). In this case, they would set the tilt of the lighting direction on the response object  $180^\circ$  from the tilt on the test object. With correct disparities, the shape of the test object was well specified and observers would therefore set the tilt of the lighting on the response object to a value close to the tilt of the lighting on the test object. Thus, the condition with zero disparities yields the same predictions for the image- and shape-based methods, and the condition with correct disparities yields entirely different predictions for the two methods.

## Methods

### Observers

Three female observers participated. They were 22–27 years of age and had normal visual acuity and stereopsis. They wore their optical corrections during testing. They were experienced psychophysical observers but were unaware of the experimental hypothesis.

### Lighting parameters

We presented eight lighting tilts ( $0, 45, 90, 135, 180, 225, 270,$  and  $315^\circ$ ) while keeping the lighting slant at  $30^\circ$ .

### Shape conditions

We presented the test objects with two different combinations of shape cues.

- A. *Shading and global concavity*: We made disparity uninformative in this condition by presenting the test objects with zero disparity. They were shaded appropriately for a globally concave object. The image- and shape-based methods both predict  $180^\circ$  tilt errors.

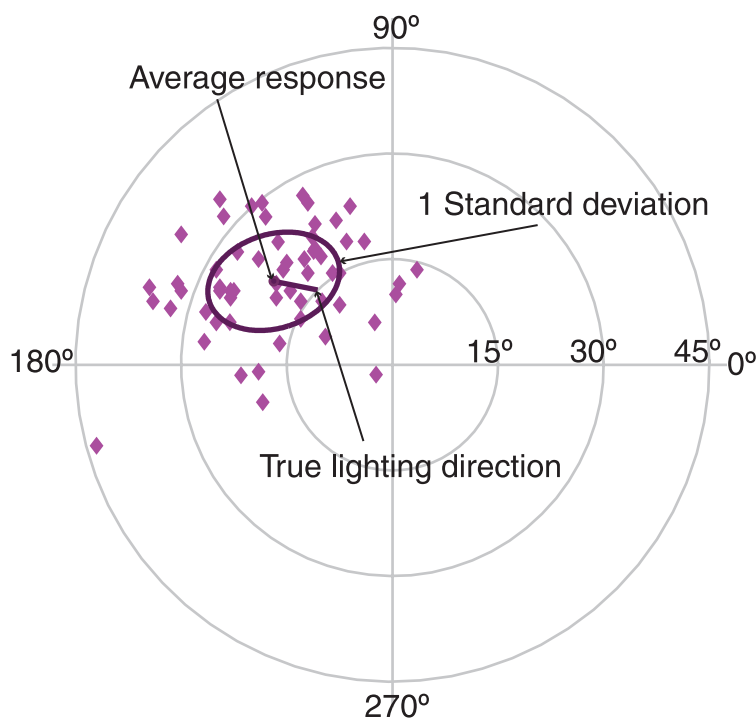


Figure 4. Responses from one subject for one lighting direction. The dots represent individual settings plotted in polar coordinates in which the angle is lighting tilt and the radius is lighting slant. The ellipse is centered on the mean of the settings and is adjusted in size and orientation to contain all of the settings out to one standard deviation from the mean. The ellipse axes are oriented along the principal components of the data, which are the orthogonal vectors that account for the maximum amount of variance (determined by a principal component analysis using the Karhunen–Loeve transform; Jolliffe, 2002; Leong & Carlile, 1998). We draw a line from the true lighting direction to the mean response to show the bias in the settings.

**B. Shading, global concavity, and binocular disparity:** The test objects again were globally concave with appropriate shading. Disparities were correct and therefore specified their true shape. Image-based methods predict 180° tilt errors, and shape-based methods predict small errors.

Before running the experiment, we familiarized observers with the physics of lighting and shading by showing sample surfaces with various lighting directions. We did not use these surfaces as stimuli in the actual experiment. We collected 20 settings for each observer for each lighting direction in each disparity condition, yielding 320 settings per observer.

## Results

Each setting is the lighting direction on the response object that the observer perceived as the same as the lighting direction on the test object. Figure 4 shows how we plot the

settings for each observer in each condition. Each dot is one setting—a combination of lighting slant and tilt—in polar coordinates. The ellipses are best fits to capture one standard deviation in all directions. The line segments connect the actual lighting direction with the average setting.

Figure 5 summarizes the individual observer and average data for the two conditions. The columns and rows show the data from different observers and different conditions, respectively. All observers behaved similarly, so we can focus on the data averaged across observers, which are shown in the rightmost column. Without reliable shape information to specify that the test surfaces are concave, all observers made large errors in Condition A, primarily due to 180° tilt errors. The average angular error was 62.0° (Figure 5), and the average tilt error was 177.6° (Figure 6). When correct disparity information reliably specified the 3D shape, observers made 5.6° average angular errors (Figure 5), and the tilt error was only 0.46° (Figure 6).

Adding the correct disparity information had a significant effect on the results. Specifically, the errors in setting

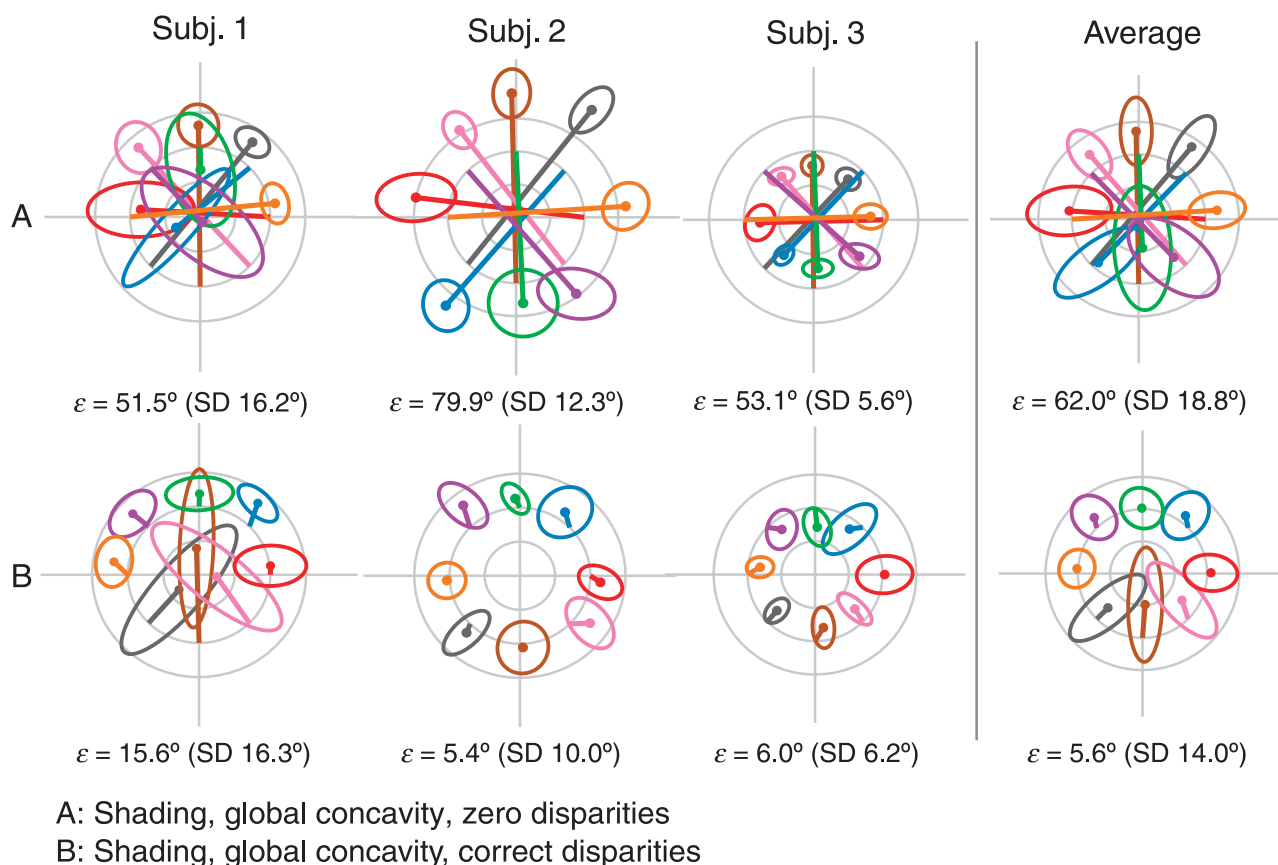


Figure 5. Summary of data from Experiment 1. (Top) In Condition A, the globally concave surfaces were specified using only shading information. The disparities were set to zero. (Bottom) In Condition B, the same surfaces were specified with shading and correct disparities. Colors represent different lighting tilts on the test object. The ellipses summarize the mean and standard deviation of the settings for each condition with lighting tilt and slant represented as in Figure 4. The line segments connect the center of each ellipse (the average setting) to the true lighting direction. The numbers under each plot are the average angular error ( $\epsilon$ ) and the average standard deviation of the settings ( $SD$ ).

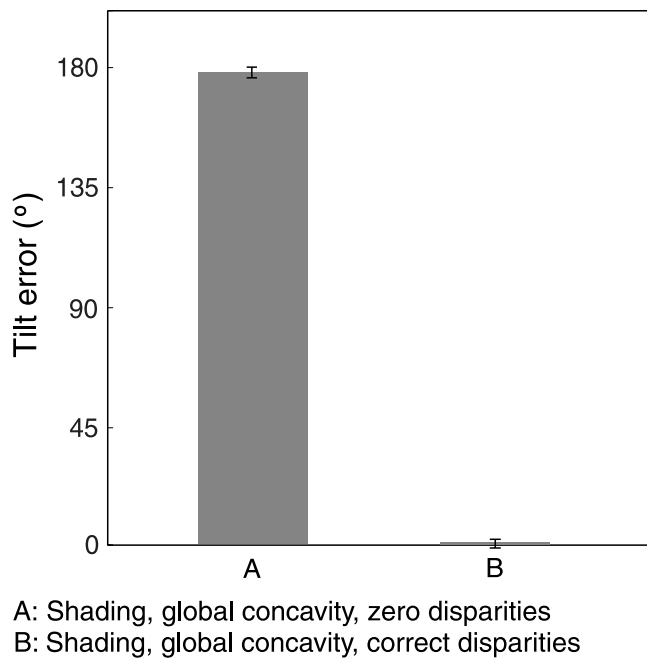


Figure 6. Average tilt errors across observers for the two conditions. We calculated the tilt error by measuring the angular tilt difference between the average response and the true lighting direction. Error bars are standard errors of the mean. Observers averaged 177.6° tilt errors when the concave shape of the test object was poorly specified (Condition A). When the shape of the test object was well specified (Condition B), observers average 0.46° tilt errors.

lighting direction were significantly smaller when the test object's shape was well specified.

## Discussion

The same stimuli were used in the two conditions, so the 2D image content was identical. Thus, any image-based method would have yielded the same pattern of settings whether disparity was informative or not. This means that observers used the 3D shape information to match the true lighting directions even though it produced opposing 2D shading patterns on the test and response objects. The results, therefore, demonstrate that people use a shape-based approach to estimate lighting direction.

## Experiment 2

We next examined how 3D shape information affects estimates of lighting direction. Specifically, we varied the shape cues used to specify the test objects in the same matching task. Our analysis of shading suggests that with reliable 3D shape information, observers should be able to

accurately estimate the lighting direction. When the 3D shape is poorly specified, we expect observers to rely more on their prior expectations of lighting direction.

## Methods

### Observers

Four female observers participated. They were 22–28 years of age and had normal visual acuity and stereopsis. They wore their optical corrections during testing. They were experienced psychophysical observers but were unaware of the experimental hypotheses.

### Lighting parameters

We presented each combination of four lighting slants (0, 15, 30, and 45°) and eight lighting tilts (0, 45, 90, 135, 180, 225, 270, and 315°). Tilt is undefined when slant is 0°, so we considered a total of 25 combinations of lighting directions.

### Shape conditions

We presented the test objects with four different combinations of shape cues (Figure 7).

- A. *All cues present*: The test objects were rendered with shading, global convexity, occluding contour, binocular disparity, and texture gradient. The cue of familiar shape was also present in that the test object was a sphere, a well-known shape. Because all cues were present in Condition A, the 3D shape was very well specified.
- B. *Shading, global convexity, and occluding contour*: We eliminated disparity by presenting the stimuli monocularly. We eliminated the texture gradient by deleting the random-element texture. We eliminated familiar shape by using 3D shapes that were randomly perturbed as shown in Figure 7. By comparing responses in Condition B to those in Condition A, we could assess the combined contribution of the texture gradient, disparity, and familiar shape in specifying 3D shape and thereby aiding the estimation of lighting direction.
- C. *Shading and global convexity*: We eliminated the occluding contour by presenting the stimulus in a circular software aperture. By comparing responses in this condition to those in Condition B, we could determine the role of occluding contour in specifying 3D shape in the estimation of lighting direction.
- D. *Shading only*: We eliminated global convexity by creating the stimulus from a frontoparallel plane (rather than a sphere) that was randomly perturbed in depth as shown in Figure 7. We clipped the stimulus with a square aperture to avoid an additional



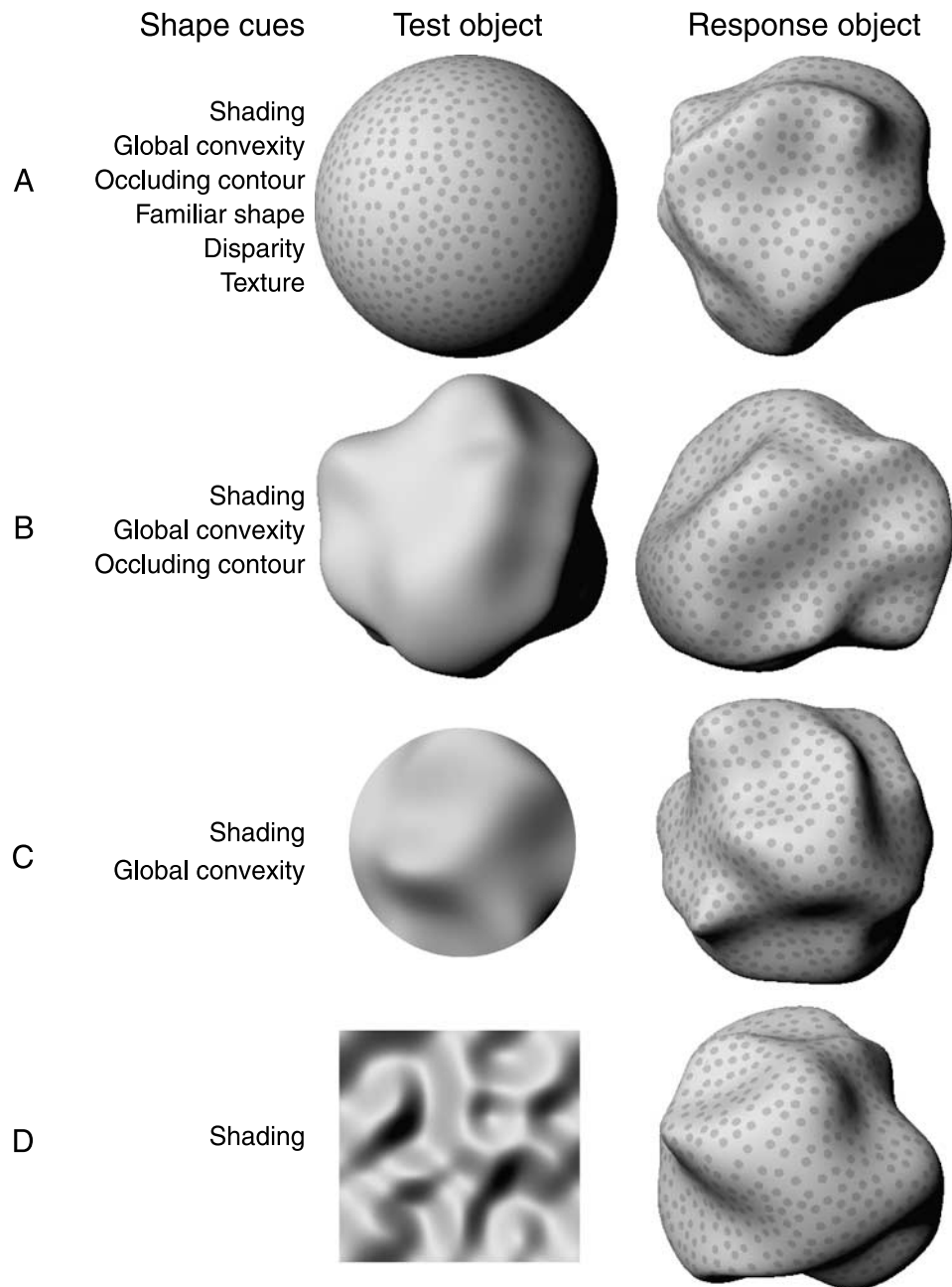


Figure 7. Sample stimuli. The shape of the test object was specified by different sets of cues in each condition (listed on the left). The shape of the response object was specified by the same set of cues in all four conditions (disparity, texture gradient, global convexity, occluding contour, and shading).

cue to convexity. By comparing responses in this condition to those in Condition C, we could determine the role of global convexity in specifying 3D shape and thereby aiding the estimation of lighting direction. Performance in this condition also tells us how well people can use shading alone to estimate light direction.

We collected 20 settings from each observer for each lighting direction in each shape condition, yielding 2000 settings per observer.

## Results

Figure 8 summarizes the individual observer and average data for the four test-object conditions. The columns and rows show the data from different observers and different conditions, respectively. The data were quite similar across observers, so we can focus on the data averaged across observers, which are shown in the rightmost column. Changing the set of available shape cues had a systematic effect on observers' settings. The left panel of Figure 9 plots the average angular difference

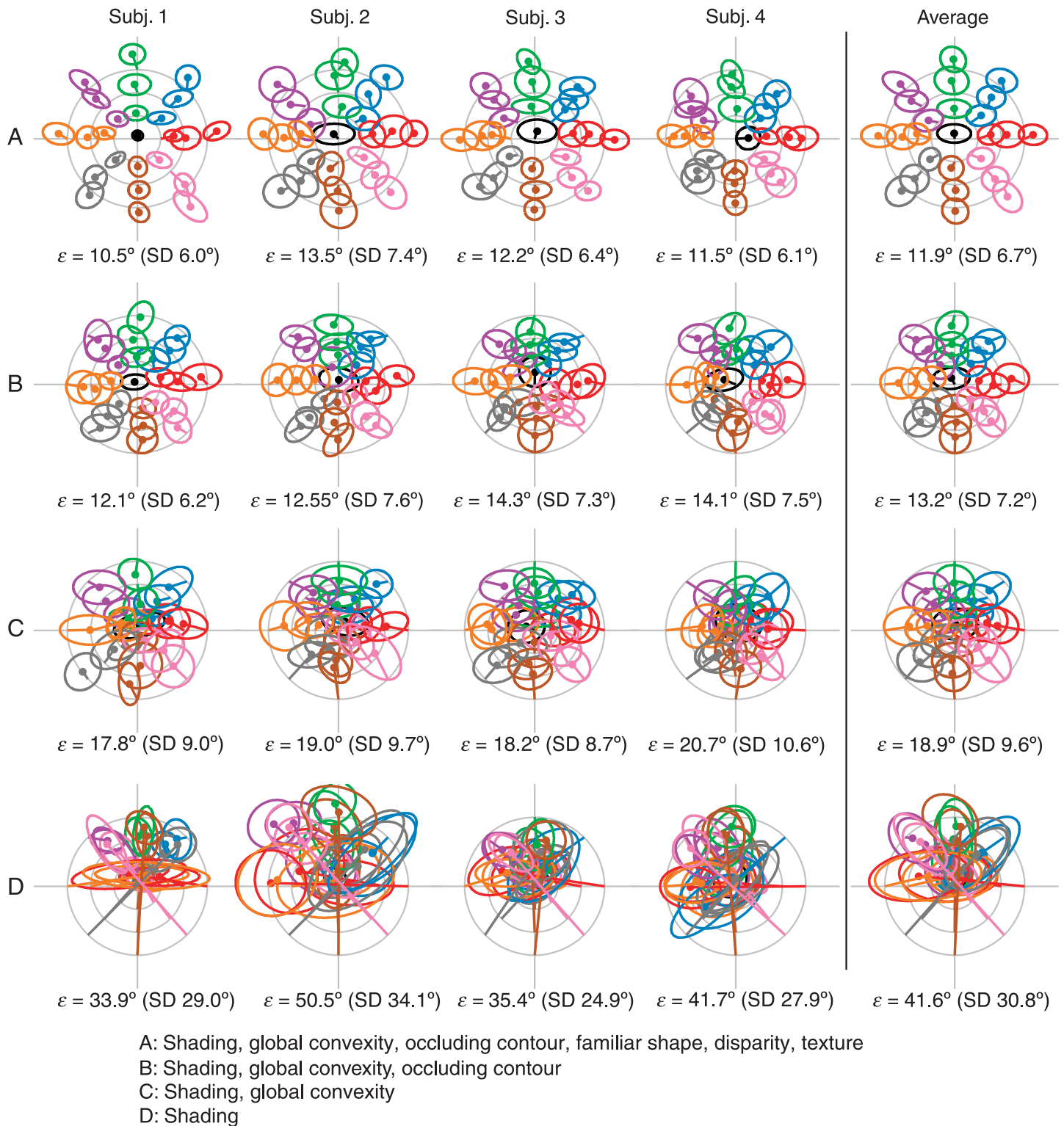


Figure 8. Summary of the data from all shape-cue conditions and observers. The columns show data from different observers and the average. The rows show the data from the four test-object conditions. Colors represent different lighting tilts on the test object. The ellipses summarize the mean and standard deviation of the settings for each condition, lighting tilt, and lighting slant as described in Figure 4. The line segments connect the average setting to the actual lighting direction. Numbers under each plot are average angular error ( $\epsilon$ ) and average standard deviation of the settings ( $SD$ ).

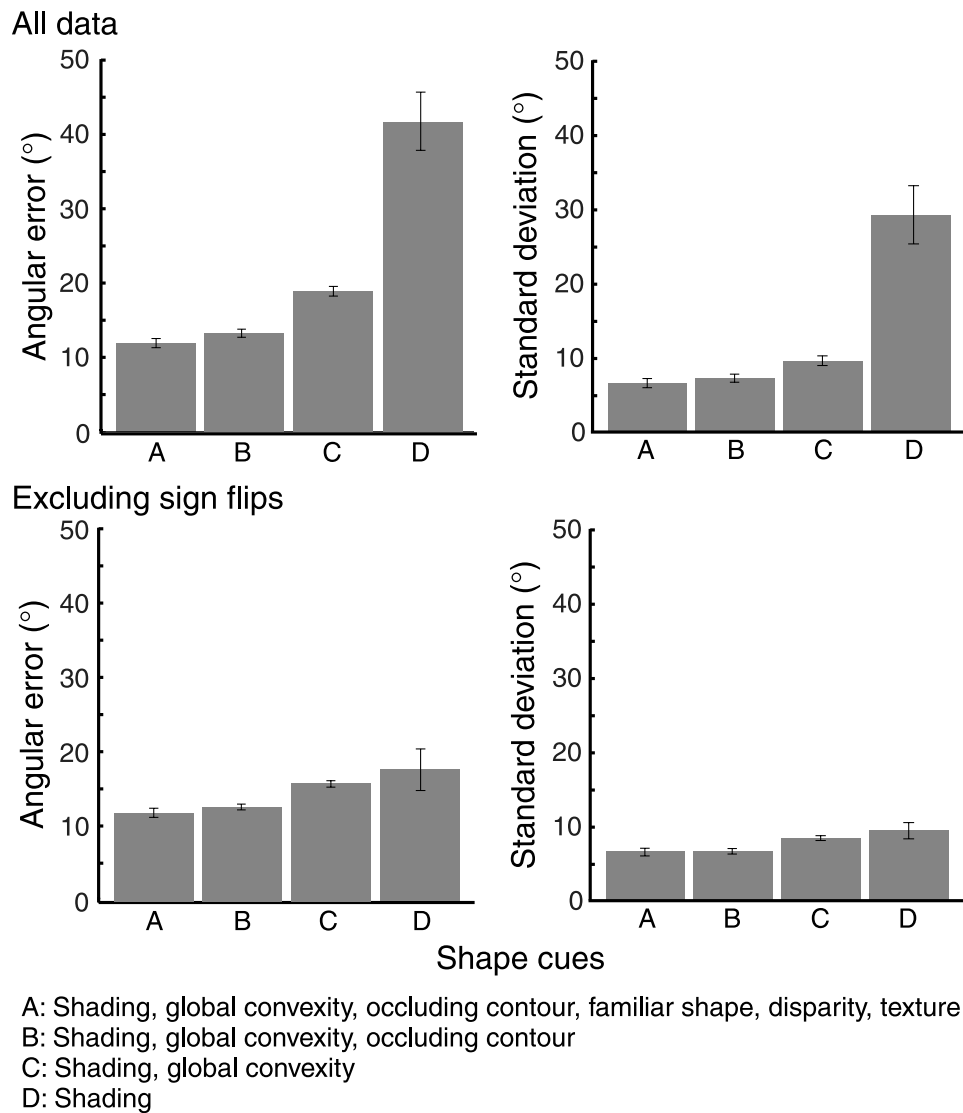


Figure 9. (Upper row) Average angular errors (left) and average standard deviations (right) for each shape-cue condition using all trials. The errors and standard deviations are expressed in angular units irrespective of the light direction. Error bars represent standard errors of the mean. (Lower row) Average angular errors (left) and average standard deviations (right) for each shape-cue condition after excluding flips in the sign of the tilt. We define a sign flip as any trial with a tilt error between  $135^\circ$  and  $225^\circ$ .

between the actual and responded lighting directions, while the right panel plots the standard deviation of the settings.

To determine which effects were statistically reliable, we conducted repeated-measures ANOVAs with angular error and with standard deviation as dependent measures. With angular error as the dependent measure, there were significant effects of shape-cue condition, lighting slant, and lighting tilt on angular error ( $p < 0.001$  in all three cases); there were also significant interactions of condition and slant and of condition and tilt ( $p < 0.001$  in both cases). With standard deviation as the dependent measure, there were significant main effects of condition, slant, and tilt ( $p < 0.001$ ) and again significant interactions of condition and slant and of condition and tilt ( $p < 0.001$ ).

Settings were most accurate in the full-cue condition (Condition A). Figure 8 shows that the angular errors in this condition were smallest and did not vary systematically with lighting tilt or slant. The average angular error was only  $11.9^\circ$  (Figure 9, upper row). The settings were also the most precise in this condition. The best-fitting ellipses in Figure 8 were small for all tilts and slants. The average standard deviation was only  $6.7^\circ$  (Figure 9, upper row).

The settings in Condition B were somewhat less accurate than those in Condition A. The average angular error and average standard deviation were slightly greater at  $13.2^\circ$  and  $7.2^\circ$ , respectively (Figure 9, upper row). These values were significantly greater than in Condition A:  $t(6) = 1.6$ ,  $p = 0.04$  (one-tailed) and  $t(6) = 1.4$ ,  $p = 0.05$

(one-tailed), respectively. The small decrease in performance shows that the cues of familiar shape, disparity, and texture provided useful information for specifying shape and thereby aided the estimation of lighting direction. It is somewhat surprising, however, that removing these shape cues had such a small effect; we will return to this observation in the [Discussion](#) section.

The settings in Condition C were less accurate than those in Condition B. The average angular error and standard deviation were now 18.9° and 9.6°, respectively ([Figure 9](#), upper row). Both of these values were significantly greater than in Condition B:  $t(6) = 6.7$ ,  $p < 0.001$  and  $t(6) = 4.5$ ,  $p < 0.01$ . The decrease in accuracy and precision means that the occluding contour (the cue not presented in Condition C) helped specify the shape of the test object and that observers used this greater specification to make better settings.

The settings in Condition D were much less accurate than those in Condition C. The average angular error and standard deviation were 41.6° and 30.8°, respectively ([Figure 9](#), upper row). Both of these values were significantly greater than in Condition C:  $t(6) = 5.7$ ,  $p < 0.001$  and  $t(6) = 10.2$ ,  $p < 0.001$ . These results show that observers are much better at determining the lighting direction when the object is globally convex than when it is not. The pattern of errors is particularly interesting. [Figure 8](#) shows that when the lighting direction was below the line of sight (i.e., lighting tilt was between 180 and 360°), observers often made tilt errors of  $\sim 180^\circ$  in their settings. In other words, they perceived the light as above the line of sight even though it was below. This pattern of responses resulted in a bimodal distribution of settings and contributed to the large angular errors in Condition D ([Figure 9](#), upper row).

This observation is clearer in [Figure 10](#), which plots average angular error as a function of lighting slant and

tilt. Notice that tilt had essentially no effect on error in Conditions A, B, and C but had a large and systematic effect in Condition D. In particular, large errors were observed when the tilt was between 180 and 360°, i.e., cases in which the actual light direction was below the line of sight. We also examined angular errors after excluding trials with tilt errors of  $\sim 180^\circ$ . Specifically, we excluded trials for which the tilt error was between 135 and 225°. The results are shown in the lower row of [Figure 9](#). The average angular error and standard deviation for Conditions A–C are nearly unchanged, but the average error in Condition D decreased from 41.6° to 17.7° and the standard deviation from 30.8° to 9.3°. Thus, the error pattern in Condition D shows that shading information alone is not sufficient for viewers to estimate lighting direction; when shape is not specified by other cues, they tend to see the light as coming from above the line of sight even when it is actually coming from below.

## Bayesian model

To further analyze the data, we used a Bayesian framework to represent the information about lighting direction contained in the sensory data and the information provided by previous experience, and the means by which observers should combine such information. Bayes’ Rule provides the optimal method (Kersten, Mamassian, & Yuille, 2004):

$$P(L|I) \propto P(I|L)P(L). \quad (8)$$

The first term on the right side of the equation is the likelihood distribution, which represents the information in the sensory data (i.e., the image  $I$ ). In this paper, we do

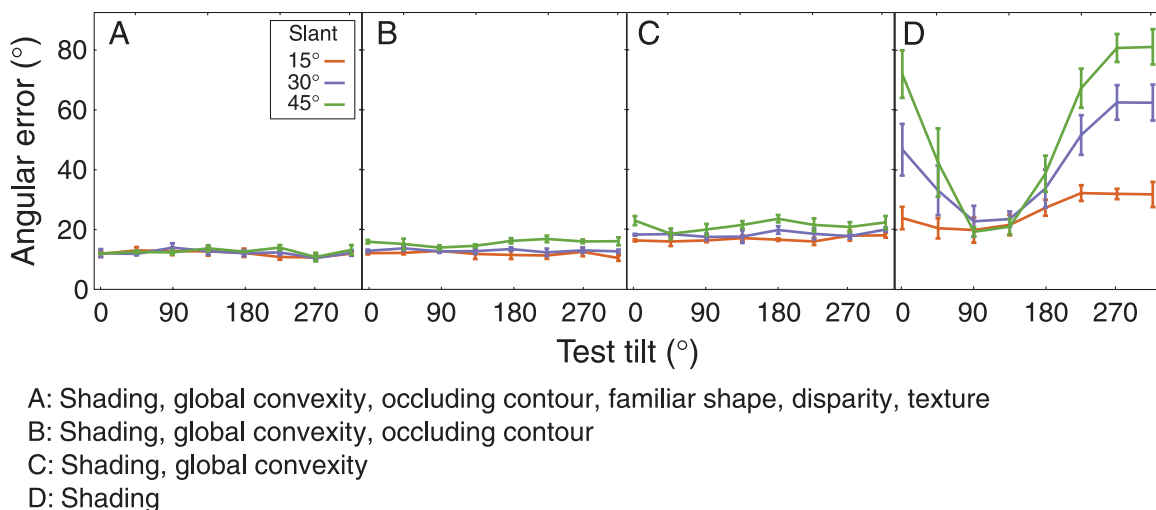


Figure 10. Angular error for different lighting tilts. Each panel plots angular error averaged across observers as a function of lighting tilt. Different colors represent the data for different lighting slants, red for 15°, blue for 30°, and green for 45°. (A–D) The data from Conditions A–D. Error bars represent standard errors of the mean.

not present a generative model of how surface shape, material properties, and illumination combine to produce the likelihood distribution. We simply use the distribution to represent the light-direction information available in the sensory data. We assume that the likelihood distribution is unbiased. The second term on the right side of the equation is the prior, which represents the distribution of likely lighting directions independent of the sensory data. We know that observers tend to assume that light comes from above and slightly to the left (Adams, Graf, & Ernst, 2004; Mamassian & Goutcher, 2001; O’Shea, Agrawala, & Banks, 2008; Sun & Perona, 1998). Observers should base their estimates of lighting direction on the product of the likelihood and prior, which is the posterior distribution on the left side of Equation 8.

We parameterized lighting directions in spherical coordinates, so we used Von Mises–Fisher (VMF) distributions to model the data. The VMF distribution is an isotropic continuous probability distribution that describes spherical data with a mean of  $\mu$  and a concentration of  $\kappa$ . The distribution on a sphere for  $\mathbf{x} \in \mathbb{R}^3$  is

$$\gamma(\mathbf{x}; \mu, \kappa) = \frac{\kappa}{4\pi \sinh(\kappa)} e^{\kappa(\mathbf{x} \cdot \mu)}, \tag{9}$$

where  $\kappa \geq 1$  and  $\|\mu\| = \|\mathbf{x}\| = 1$ . The parameter  $\mu$  has the coordinates  $[\cos(\theta)\sin(\varphi), \sin(\theta)\sin(\varphi), \cos(\varphi)]$ , which

correspond to the Cartesian coordinates of lighting slant ( $0^\circ \leq \varphi \leq 90^\circ$ ) and tilt ( $0^\circ \leq \theta \leq 360^\circ$ ). The parameter  $\kappa$  is inversely proportional to the spread of the distribution, so as  $\kappa$  increases, the variance of the distribution decreases. Figure 11 shows some sample distributions.

We assumed that observers based their judgments on the peak of the posterior distribution, which is proportional to the product of the likelihood and prior (Equation 8). In particular, we assumed that the judgments were derived from the maximum of the posterior. We then found the likelihood and prior distributions that best predicted the observers’ responses for each experimental condition. In doing so, we assumed that the likelihood distributions were unbiased (that is, that the peaks of those distributions were centered on the true lighting direction). In finding the best-predicting distributions, the likelihoods had one free parameter  $\kappa$  for each of the four test-object conditions. Thus, we found  $\kappa_A$  for the data in Condition A, and likewise  $\kappa_B, \kappa_C,$  and  $\kappa_D$  for the appropriate data sets. The prior had two free parameters for the coordinates of the peak of the distribution ( $\varphi_P$  and  $\theta_P$ ) and one parameter  $\kappa_P$  for the spread. We found one set of prior parameters for all four conditions because we assumed an observer’s prior did not change over the course of the experiment.

Because we assumed unbiased likelihoods, we set the means of the likelihoods equal to the coordinates of the

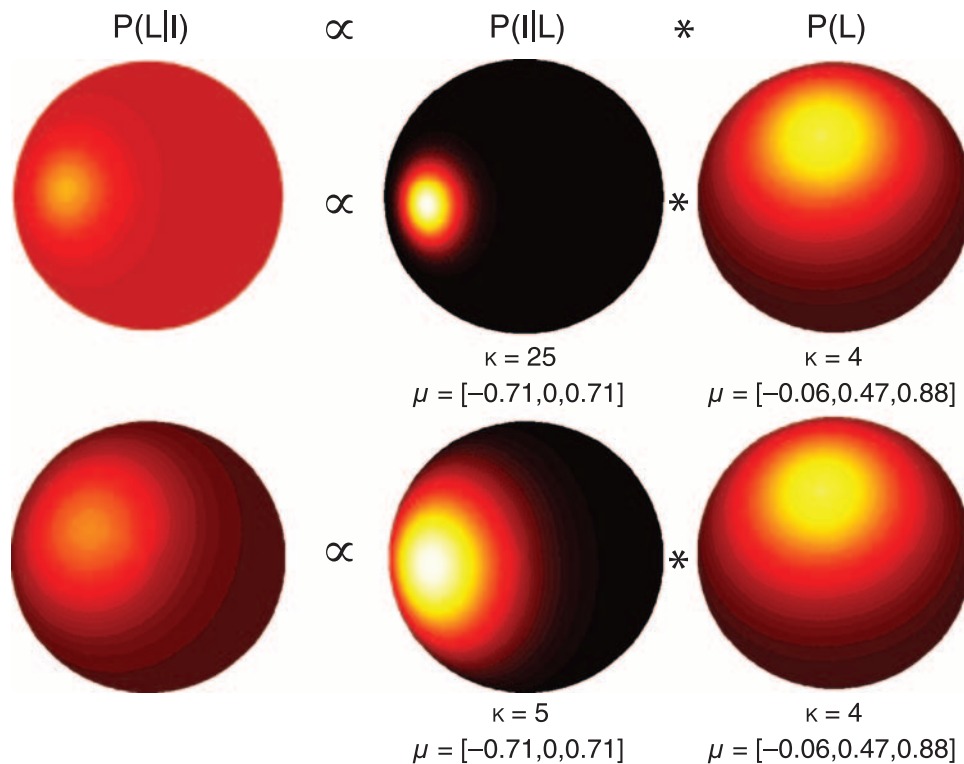


Figure 11. Von Mises–Fisher (VMF) distributions and implementation in Bayesian framework. (Upper row) From left to right, the VMF distributions represent the posterior, likelihood, and prior distributions in a Bayesian framework. The posterior is the product of the likelihood and prior. Here the likelihood has lower variance than the prior, so the posterior is similar to the likelihood. (Lower row) A likelihood with greater variance (lower  $\kappa$ ) leads to a greater influence of the prior on the posterior.

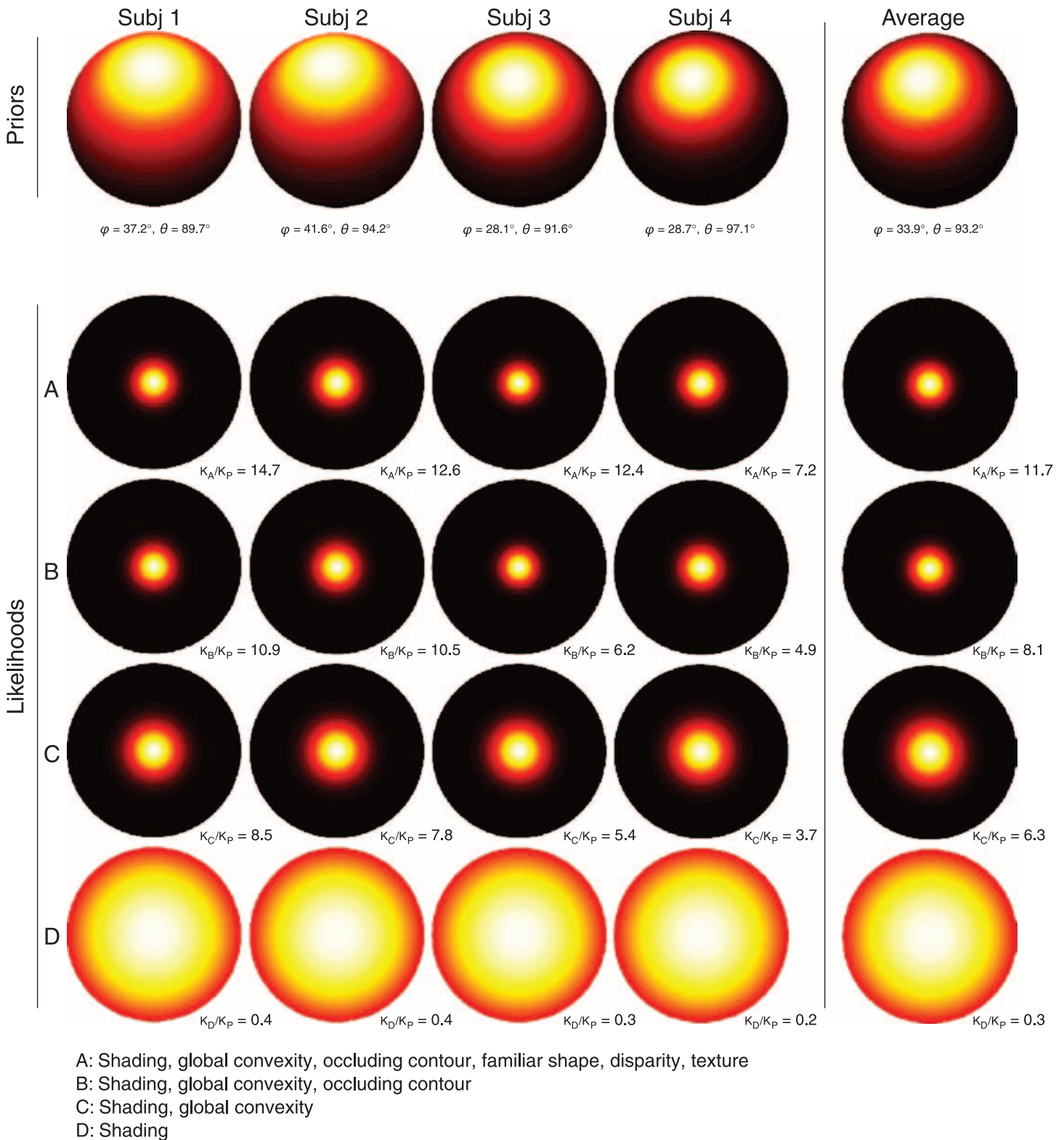


Figure 12. The likelihood and prior distributions that best fit each observer's data. The first four columns represent the distributions for individual observers. The rightmost column represents the distributions derived from fitting the combined data from all four observers. The ratio of the best-fitting likelihood variance to the best-fitting prior variance is given below each likelihood panel. In this graphical representation, the means of the likelihoods have been set to  $[\varphi = 0, \theta = 0]$ , but the means were actually the true lighting slants and tilts for each lighting direction.

actual lighting direction in each condition. As we said, there were four parameters for the variances of the likelihoods ( $\kappa_A$ ,  $\kappa_B$ ,  $\kappa_C$ , and  $\kappa_D$ ) and one for the variance of the prior ( $\kappa_P$ ). However, the position of the maximum of the posterior is determined by ratios of likelihood  $\kappa$  and prior  $\kappa$ , so there were only four free parameters for  $\kappa$ . To deal with this constraint, we set  $\kappa_D$  to 1 and found the best values for the other four. Thus, we fixed the likelihood locations for all lighting directions within a shape condition. We found the best values for the six free parameters for the complete set of data from each observer using a non-linear, least-squares optimization routine (Matlab’s `lsqnonlin` routine). The routine found the set of parameters that minimized Chi square ( $\chi^2$ ), the sum of the squares of the angular errors. We did not attempt to fit the variances of the observer settings.

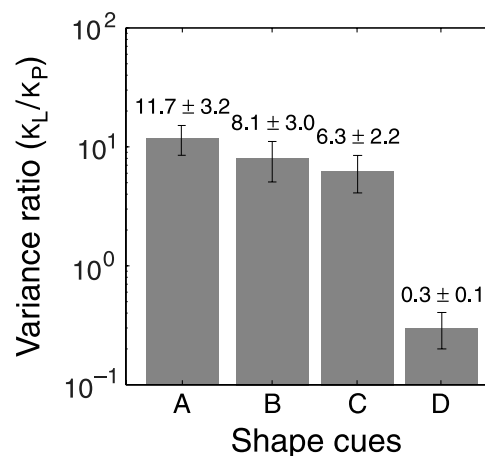
Figure 12 displays the likelihood and prior distributions that best fit each observer’s data. The first four columns represent the results for the four observers and the rightmost column the results averaged across observers. The top row represents the best-fitting prior distributions and the next four rows the best-fitting likelihood distributions for Conditions A, B, C, and D, respectively. Note that the means of the likelihoods have been plotted at [0, 0] because many different directions were actually presented and could not be readily shown in one graph.

The results were strikingly similar across observers. For example, the prior distribution is centered above the visual axis for all four observers; specifically, the best-fitting tilt varies from 89.7 to 97.1°; tilts greater than 90° are counterclockwise from vertical. This result is consistent with the aforementioned light-from-above prior (Adams

et al., 2004; Mamassian & Goutcher, 2001; O’Shea et al., 2008; Sun & Perona, 1998). The prior distribution is also roughly equally displaced from the origin in all four observers; the best-fitting slant varies from 28.1 to 41.6°. This result is nicely consistent with our earlier finding that the assumed slant for lighting direction is 20–30° above the line of sight (O’Shea et al., 2008). The best-fitting likelihood distributions were also remarkably similar across observers. The spread of the distributions increased in quite similar fashion for all four observers as we took shape information away in going from Condition A to Condition D.

As we said earlier, the location of the maximum of the product of two VMF distributions is determined by the ratio of the distributions’ variances. The ratio reflects the degree to which the likelihood or prior determines the location of the posterior. In Figure 13, we plot the average ratio of the likelihood and prior variances—e.g.,  $\kappa_L/\kappa_P$ —for all observers for each shape condition. The ratio is large in Condition A where all shape cues were present, which is consistent with the fact that observers made quite accurate settings in that condition. As shape cues were taken away, the ratio became smaller, which is consistent with the observation that observers made successively less accurate settings as their settings drifted toward above the line of sight. Indeed, the ratio is less than 1 for Condition D where only shading was available, consistent with the observers relying primarily on their prior expectation of lighting direction in that case.

We next investigated how well our model fit the data compared with other plausible models. To do this, we computed  $\chi^2$  for four models.



- A: Shading, global convexity, occluding contour, familiar shape, disparity, texture
- B: Shading, global convexity, occluding contour
- C: Shading, global convexity
- D: Shading

Figure 13. The ratio of the likelihood  $\kappa_L$  and prior  $\kappa_P$  for the four shape conditions. The ordinate is the likelihood variance parameter (e.g.,  $\kappa_A$ ,  $\kappa_B$ ) divided by the prior variance parameter ( $\kappa_P$ ). The values that went into computing the ratio are the across-observer averages. Error bars represent the standard error of the mean.

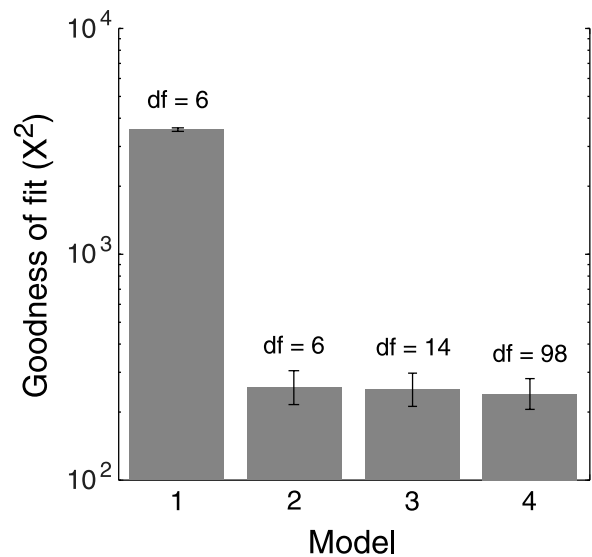
1. The first was a random model with six free parameters (the same six as in the model that generated Figure 12). In this model, we first randomly reassigned settings to conditions (with replacement) and then we fit the parameters to the data. This model provides an upper bound on our measure of goodness of fit for comparison with the fits of the other models.
2. The second was the model described earlier that was used to generate Figure 12. There are six free parameters in this model.
3. The third model was similar to the second except that the likelihood variance parameters ( $\kappa_A$ ,  $\kappa_B$ ,  $\kappa_C$ , and  $\kappa_D$ ) were allowed to differ for each of the four lighting slants. The parameters of the prior were the same as in Models 1 and 2. Thus, this model has 14 free parameters (two for the prior and 12 for the likelihoods).
4. The fourth model was similar to the third except that the likelihood variance parameters were allowed to differ for each combination of lighting slant and tilt. The parameters of the prior were the same as in the above models. Thus, this model has 98 free parameters (two for the prior and 96 for the likelihoods).

Figure 14 shows goodness of fit ( $\chi^2$ ) for the four tested models. As one would expect, the random model provides a poor fit. The other three models provide roughly equivalent fits. Models 3 and 4 have many more free parameters than Model 2, yet Model 2 provides essentially the same fit to the data. We conclude that Model 2—the one used to generate Figure 12—provides the most parsimonious account for the data.

## Discussion

### Summary of results

In Experiment 1, we showed that observers use 3D shape information to match the lighting direction in a scene. In Experiment 2, we examined how specific shape cues affect observer estimates of lighting direction. Our results show, as expected, that accurate perception of lighting direction depends on reliable shape information. When the 3D shape of the object was specified by many robust shape cues, observers estimated direction accurately. When the shape was poorly specified, responses were very inaccurate: In that case, the perceived direction was above the view direction and slightly counterclockwise from vertical no matter what the actual light direction was. We used a Bayesian framework to model the data. The framework combined light-direction information contained in the images with a light-direction prior. The prior was centered above and slightly to the left: tilt and slant of  $93.2^\circ$  and  $33.9^\circ$ , respectively.



- 1: Random model  
 2: Likelihoods do not vary  
 3: Likelihoods vary by slant  
 4: Likelihoods vary by slant and tilt

Figure 14. Goodness of fit for the models tested. Chi square ( $\chi^2$ ) is plotted for each model. Model 1 is a random model in which the observed settings were randomly reassigned to condition and we then found the six parameters that provided the best fit to the reassigned data. Model 2 is the model used to generate Figure 12; we found the six parameters that provided the best fit to the observers’ settings across experimental conditions. Model 3 is similar to Model 2 except that the likelihood  $\kappa$  is allowed to differ for different lighting slants; the model has 14 free parameters. Model 4 is similar except the likelihood  $\kappa$  is allowed to differ for different lighting slants and tilts; it has 98 free parameters.

We summarize the findings with a simple demonstration in Figure 15. The upper panel is a shaded image of a surface whose 3D shape is poorly specified. The surface is globally flat, the occluding contour is not visible, and disparity and texture are not available; the shape is specified by shading only. Notice that the light source appears to be above the panel. The lower panel is the same shaded image, but now 3D shape is well specified by disparity and the texture gradient. It is now evident that the light source is actually below. The figure shows that the light direction is correctly perceived when the shaded object’s 3D shape is well specified and is incorrectly perceived to be in the direction of the light-from-above prior when we specify the 3D shape using only shading.

### Light-direction prior

There is a great deal of evidence that viewers assume light comes from above and slightly to the left. Convexity–concavity judgments are consistent with an assumed



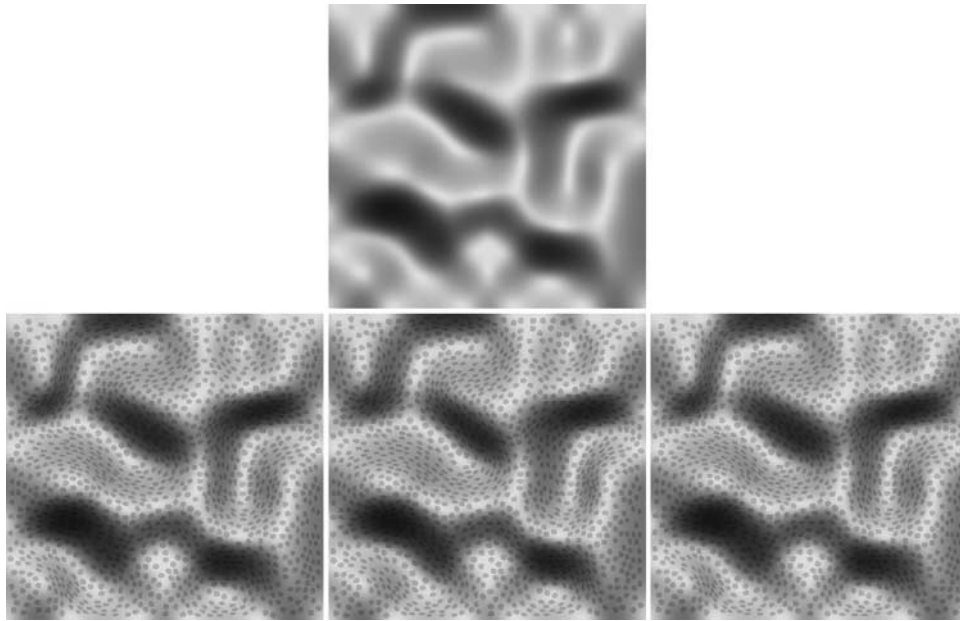


Figure 15. Demonstration of how shape information affects the estimation of lighting direction. (Upper row) A shaded object in which the 3D shape is poorly specified. Most viewers of this image think the illumination is from above. (Lower row) The same object and illuminant, but now the shape is well specified. For cross-fusing, use the two panels on the left. For divergent fusing, use the two on the right. When the image is fused, the object's 3D shape is specified by disparity and texture as well as shading. If you are correctly fusing, you should see the brighter parts of the surface as slanted top near and bottom far. Most viewers of the lower panels now think the illumination is from below. The reliable shape cues in the lower panels allow one to estimate the true light direction (Equation 4). Those cues are not available in the upper panel, so the light-direction prior dictates the perceived light direction.

lighting tilt of  $\sim 110^\circ$  (Adams et al., 2004; Jenkin, Jenkin, Dyde, & Harris, 2004; Mamassian & Goutcher, 2001; Morgenstern & Murray, 2009; Sun & Perona, 1998). Speed of visual search is greatest when the tilt is roughly the same value (Adams, 2007; Enns & Rensink, 1990; Kleffner & Ramachandran, 1992). To our knowledge, only one previous study has made measurements relevant to specifying the slant of the lighting prior. O'Shea et al. (2008) showed that 3D shape judgments of shaded objects were most accurate when the lighting slant was  $20\text{--}30^\circ$ .

With the exception of Morgenstern and Murray (2009), our task was very different from the ones used in the above-mentioned studies. We estimated the prior for light direction by having observers adjust the direction of the illumination on an object whose shape was well specified to match the perceived direction of the illumination on an object whose shape was poorly specified. The average tilt of the prior was  $93.2^\circ$  and the average slant was  $33.9^\circ$ . These estimates of the prior parameters are remarkably consistent with the estimates from previous work despite the use of an entirely different task.

### Perceiving lighting inconsistencies

Ostrovsky, Cavanagh, and Sinha (2005) reported that people have considerable difficulty detecting inconsistencies in the direction of lighting in scenes composed of

several objects. In a visual search task, they presented nine objects. In one condition, all nine were illuminated with the same light. In another, eight of the nine were illuminated with one light and one was illuminated with a light whose tilt differed by  $90^\circ$ . The task was to indicate whether the lighting was consistent or inconsistent. The shapes were reasonably well specified; the stimuli were most similar to the test objects of Condition B in our experiment. Ostrovsky et al. found that people could discriminate the inconsistent from the consistent displays, but performance was far from perfect. The relatively poor performance seems inconsistent with our results. In our Condition B, observers made accurate and precise settings. Average angular error was  $13.2^\circ$  and the standard deviation was  $7.2^\circ$ , values that are much lower than the  $90^\circ$  differences in the lights in the experiment of Ostrovsky et al. Their result is similar to findings that people have difficulty detecting inconsistencies in attached and cast shadows in complex scenes (Farid & Bravo, 2010; Mamassian, 2004).

Ostrovsky et al. hypothesize that the visual system can compute illumination direction for individual objects when shape is well specified and this is consistent with our data. They also speculate that multiple estimates of light direction from various objects "may not support any accumulation into a group direction" (p. 1311). Thus, the limit may have to do with accumulating estimates from individual objects into one global estimate of scene illumination.

## Effectiveness of different shape cues

We observed a small decrease in performance between Conditions A and B. Thus, removing the robust shape cues of disparity, texture, and familiar shape had a small but noticeable effect. This result means that observers could use the cues of occluding contour, global convexity, and shading to estimate lighting direction reasonably accurately, which is consistent with previous work on illumination matching (Pont & Koenderink, 2007). We also found a small decrease in performance between Conditions B and C, which means that removing the occluding contour had a discernible but small effect. This result suggests that observers could use the remaining cues of global convexity and shading to estimate light direction fairly accurately. There was a large decrease in performance between Conditions C and D, which suggests that global convexity had a significant effect on the ability to estimate light direction.

We first consider the information provided by the occluding contour. As we noted earlier, the variation in luminance near the occluding contour of an object can be used to estimate the tilt of the lighting (Nillius & Eklundh, 2001; Figure 1). This estimation technique has been utilized effectively to detect illumination inconsistencies within photographs (Johnson & Farid, 2005). One cannot, however, estimate the slant of the lighting from this cue without inferring the shape of the rest of the object. Thus, the cues of global convexity and shading must have been the primary determinants of direction estimates.

We next examine the lighting information available with globally convex stimuli and relate that information to observed performance. People tend to assume that surfaces are globally convex (Langer & Bühlhoff, 2001; Mamassian & Landy, 1998). This assumption is consistent with the test objects presented in Conditions A–C. As we said earlier, lighting direction can be recovered if the object's shape is known and the surface albedo is constant. The lighting information contained in shaded globally convex objects is illustrated in Figure 16. The stimulus is the irregular object in Figure 1 seen through an aperture so that the occluding contour is not visible. This stimulus corresponds to Condition C in our experiment. In constructing the luminance maps in the upper row, we assumed that the object's shape was estimated accurately. Thus, the plots of luminance as a function of surface tilt and slant are regular with clear peaks at the slant and tilt values corresponding to the slant and tilt of the light source. The luminance maps in the lower row were constructed with the same stimulus and lighting directions, but we assumed that the object is a sphere. By making this assumption, the observer can estimate the slant and tilt of each point in the image based on the coordinates of the point in the image. The luminance maps are of course less regular than in the upper row, but they still contain the same general pattern. Is there sufficient information to make a reasonably accurate estimate of

light direction? We examined this by using the least-squares approach summarized by Equation 4.

We ran the analysis for each of the test stimuli from Condition C using two assumed sphere sizes. In the first analysis, the radius of the assumed sphere was equal to the average radius of the test shapes. The average angular error (angular difference between the estimated and true lighting directions) was  $15.8^\circ$  ( $SD = 10.4^\circ$ ). The blue stars represent the estimates. In the second analysis, the radius of the assumed sphere was equal to the radius of the aperture. This assumed shape is less consistent with the true 3D surface geometry, and the errors of the resulting estimates were slightly higher. The average angular error was  $17.4^\circ$  ( $SD = 10.5^\circ$ ). We also ran the analysis using stimuli rendered with a constant ambient light term. The ambient term changes the luminance values in the image, but the overall pattern remains the same and the resulting estimates were similar to the previous analyses.

We observed that the least-squares estimates were reasonably similar to observers' settings for most of the stimuli. (The left panel represents one of the most accurate cases and the right panel one of the least accurate.) The similarity shows that human viewers could use a shape assumption—in this case, an assumption of sphericity—to estimate light direction reasonably accurately when the actual stimulus is only globally consistent with the assumed shape.

It is also important to consider the illumination information when the stimulus is globally flat, as it was in Condition D of our experiments. If the stimulus is a plane, the luminance map is flat: The angle between surface normals and a distant light source is constant, so the luminance is constant (Equation 4). Therefore, even a correct assumption about object shape (globally or locally correct) would not allow the observer to estimate light direction. Thus, an observer prior for lighting direction should dictate their responses as we observed in Condition D.

## Applications

Our findings have implications for the construction of shaded images. In the absence of robust shape cues, such as texture, disparity, convexity, and familiar shape, our observers misestimated the direction of lighting with shaded images. When only shading was available, the Bayesian prior for light direction—slant of roughly  $30^\circ$  and tilt of slightly more than  $90^\circ$  (i.e., above and slightly to the left)—dictated the estimate. Because of the interaction between light direction and 3D shape (Equation 1), misestimating the light direction can lead to misestimating the shape. Thus, to assure reasonably accurate perception of the 3D shape of shaded images, it is important to place directional lights near the prior. Such placement is relevant to recent work on automatic lighting design (Gumhold, 2002; Lee, Hao, & Varshney, 2006; Shacked

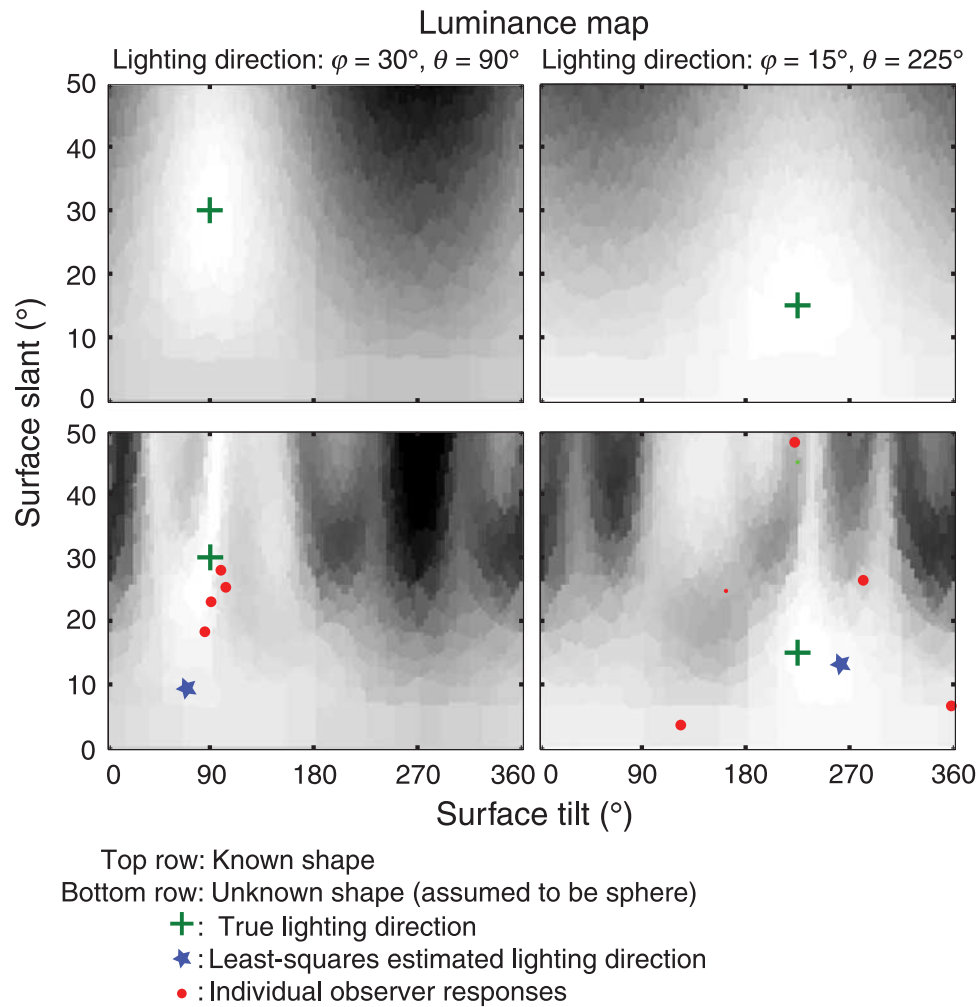


Figure 16. Luminance maps for estimated surface orientation of stimuli. The upper row shows the luminance maps for a stimulus in Condition C under the (unlikely) assumption that the 3D surface orientation was estimated accurately. The lighting direction was [ $\varphi = 30^\circ, \theta = 90^\circ$ ] on the left and [ $\varphi = 15^\circ, \theta = 225^\circ$ ] on the right. The green crosses indicate the true lighting direction. The lower row shows the luminance maps for the same stimuli under the assumption that the assumed shape is a sphere. Now the luminance map is distorted because the assumed surface orientation at each stimulus point is not necessarily correct. The green crosses again represent the true lighting directions. The blue stars represent the lighting directions estimated by the least-squares algorithm described in the [Introduction](#) section. Although the assumed shape is incorrect, the lighting-direction estimate is within  $20^\circ$  of the true direction. The red dots represent the settings from each of the four observers when presented these stimuli.

& Lischinski, 2001) and on non-photorealistic rendering techniques designed to affect the perception of 3D shape (Rusinkiewicz, Burns, & DeCarlo, 2006).

## Conclusion

In [Experiment 1](#), we found that observers use a shape-based rather than image-based approach to estimate the lighting direction of a scene. Analyses of the information contained in shading reveals that the lighting direction could in principle be correctly inferred if the reflectance properties of the surface material are known and the 3D

shape of the object generating the image is known. Our results from [Experiment 2](#) confirm this expectation. We found that observers can match the lighting directions on two objects when the shapes of the objects are well specified. We found that they set the lighting direction quite inaccurately when the shape of one of the objects is specified by shading only; instead they set the lighting direction consistent with a light-from-above prior. Thus, shading alone does not provide sufficient shape information to estimate light direction accurately. We also found that global convexity is a very effective cue in determining light direction and that this finding is expected when one considers the information contained in a globally convex object. Given our results, algorithms for producing

images with shaded objects should be consistent with viewer assumptions that the light comes from above and that most objects are globally convex.

## Acknowledgments

This work was supported by NIH Research Grant R01-EY012851 and NSF Research Grant BCS-0617701. We thank Hany Farid for discussions on the image-based approach, James O’Brien for discussions on the global convexity analysis, Yaniv Morgenstern for discussing his work on lighting cues and spherical statistics, and David Hoffman for help with fitting the model. Some of the data were previously presented at the Vision Sciences Society Annual Meeting in 2009.

Commercial relationships: none.

Corresponding author: James P. O’Shea.

Email: joshea@cs.berkeley.edu.

Address: 360 Minor Hall, Berkeley, CA 94720, USA.

## References

- Adams, W. J. (2007). A common light-prior for visual search, shape, and reflectance judgments. *Journal of Vision*, 7(11):11, 1–7, <http://www.journalofvision.org/content/7/11/11>, doi:10.1167/7.11.11. [PubMed] [Article]
- Adams, W. J., Graf, E. W., & Ernst, M. O. (2004). Experience can change the “light-from-above” prior. *Nature Neuroscience*, 7, 1057–1058. [PubMed]
- Backus, B. T., Banks, M. S., van Ee, R., & Crowell, J. A. (1999). Horizontal and vertical disparity, eye position, and stereoscopic slant perception. *Vision Research*, 39, 1143–1170. [PubMed]
- Catmull, E., & Clark, J. (1978). Recursively generated B-spline surfaces on arbitrary topological meshes. *Computer-Aided Design*, 10, 350–355.
- Dunbar, D., & Humphreys, G. (2006). A spatial data structure for fast Poisson-disk sample generation. *ACM Transactions on Graphics (Proceedings SIGGRAPH 2006)*, 25, 503–508.
- Enns, J. T., & Rensink, R. A. (1990). Influence of scene-based properties on visual search. *Science*, 247, 721–723. [PubMed]
- Farid, H., & Bravo, M. J. (2010). Image forensic analyses that elude the human visual system. *Proceedings of the SPIE*, 7541, 1–10.
- Gumhold, S. (2002). Maximum entropy light source placement. *Proceedings of the IEEE Computer Society Conference on Visualization*, 2, 275–282.
- Ho, Y. X., Landy, M. S., & Maloney, L. T. (2006). How direction of illumination affects visually perceived surface roughness. *Journal of Vision*, 6(5):8, 634–648, <http://www.journalofvision.org/content/6/5/8>, doi:10.1167/6.5.8. [PubMed] [Article]
- Ikeuchi, K., & Horn, B. K. P. (1981). Numerical shape from shading and occluding boundaries. *Artificial Intelligence*, 17, 141–184.
- Jenkin, H. L., Jenkin, M. R., Dyde, R. T., & Harris, L. R. (2004). Shape-from-shading depends on visual, gravitational, and body-orientation cues. *Perception*, 33, 1453–1461. [PubMed]
- Johnson, M. K., & Farid, H. (2005). Exposing digital forgeries by detecting inconsistencies in lighting. In *Proceedings of the 7th Workshop on Multimedia and Security (New York, NY, USA, August 01–02, 2005)* (pp. 1–10). New York, NY: ACM.
- Jolliffe, I. T. (2002). *Principal component analysis*. New York: Springer-Verlag.
- Kajiya, J. (1986). The rendering equation. *Computer Graphics*, 20, 143–149.
- Kersten, D., Mamassian, P., & Yuille, A. (2004). Object perception as Bayesian inference. *Annual Reviews of Psychology*, 55, 271–304. [PubMed]
- Kleffner, D. A., & Ramachandran, V. S. (1992). On the perception of shape from shading. *Perception & Psychophysics*, 52, 18–36. [PubMed]
- Koenderink, J. J., Pont, S. C., van Doorn, A. J., Kappers, A. M. L., & Todd, J. T. (2007). The visual light field. *Perception*, 36, 1595–1610. [PubMed]
- Koenderink, J. J., van Doorn, A. J., Christou, C., & Lappin, J. S. (1996). Perturbation study of shading in pictures. *Perception*, 25, 1009–1026. [PubMed]
- Koenderink, J. J., van Doorn, A. J., & Pont, S. C. (2004). Light direction from shaded random Gaussian surfaces. *Perception*, 33, 1405–1420. [PubMed]
- Langer, M. S., & Bülthoff, H. H. (2001). A prior for global convexity in local shape-from-shading. *Perception*, 30, 403–410. [PubMed]
- Lee, C. H., Hao, X., & Varshney, A. (2006). Geometry-dependent lighting. *IEEE Transactions on Visualization and Computer Graphics*, 12, 197–207.
- Lee, C. H., & Rosenfeld, A. (1985). Improved methods for estimating shape from shading using the light source coordinate system. *Artificial Intelligence*, 26, 125–143.
- Leong, P., & Carlile, S. (1998). Methods for spherical data analysis and visualization. *Journal of Neuroscience Methods*, 80, 191–200. [PubMed]
- Lopez-Moreno, J., Hadap, S., Reinhard, E., & Gutierrez, D. (2009). Light source detection in photographs.

- Congreso Espanol de Informatica Grafica, Sept. 9–11, 2009* (pp. 161–168). San Sebastian: CEIG.
- Malik, J., & Maydan, D. (1989). Recovering three-dimensional shape from a single image of curved objects. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *11*, 555–566.
- Mamassian, P. (2004). Impossible shadows and the shadow correspondence problem. *Perception*, *33*, 1279–1290. [[PubMed](#)]
- Mamassian, P., & Goutcher, R. (2001). Prior knowledge on the illumination position. *Cognition*, *81*, 1–9. [[PubMed](#)]
- Mamassian, P., & Landy, M. S. (1998). Observer biases in the 3D interpretation of line drawings. *Vision Research*, *38*, 2817–2832. [[PubMed](#)]
- Morgenstern, Y., & Murray, R. (2009). Contextual lighting cues can override the light-from-above prior [[Abstract](#)]. *Journal of Vision*, *9*(8):65, 65a, <http://www.journalofvision.org/content/9/8/65>, doi:10.1167/9.8.65.
- Nillius, P., & Eklundh, J. O. (2001). Automatic estimation of the projected light source direction. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *1*, 1076–1083.
- O'Shea, J. P., Banks, M. S., & Agrawala, M. (2008). The assumed light direction for perceiving shape from shading. In *Proceedings of the 5th Symposium on Applied Perception in Graphics and Visualization (Los Angeles, California, August 09–10, 2008)* (pp. 135–142). New York, NY: ACM.
- Ostrovsky, Y., Cavanagh, P., & Sinha, P. (2005). Perceiving illumination inconsistencies in scenes. *Perception*, *34*, 1301–1314. [[PubMed](#)]
- Pentland, A. P. (1982). Finding the illuminant direction. *Journal of the Optical Society of America*, *72*, 448–455.
- Pont, S. C., & Koenderink, J. J. (2007). Matching illumination of solid objects. *Perception & Psychophysics*, *69*, 459–468. [[PubMed](#)]
- Rusinkiewicz, S., Burns, M., & DeCarlo, D. (2006). Exaggerated shading for depicting shape and detail. *ACM Transactions on Graphics (Proceedings SIGGRAPH)*, *25*, 1199–1205.
- Shacked, R., & Lischinski, D. (2001). Automatic lighting design using a perceptual quality metric. *Computer Graphics Forum*, *20*, 215–226.
- Stevens, K. A. (1983). Slant-tilt: The visual encoding of surface orientation. *Biological Cybernetics*, *46*, 183–195. [[PubMed](#)]
- Sun, J., & Perona, P. (1998). Where is the sun? *Nature Neuroscience*, *1*, 183–184. [[PubMed](#)]