Two operational modes in the perception of shape from shading revealed by the effects of edge information in slant settings

Peng Sun

School of Psychology, University of Birmingham, Birmingham, UK, & Department of Cognitive Science, University of California Irvine, Irvine, CA, USA

Andrew J. Schofield

School of Psychology, University of Birmingham, Birmingham, UK

The perception of shape from shading (SFS) has been an active research topic for more than two decades, yet its quantitative description remains poorly specified. One obstacle is the variability typically found between observers during SFS tasks. In this study, we take a different view of these inconsistencies, attributing them to uncertainties associated with human SFS. By identifying these uncertainties, we are able to probe the underlying computation behind SFS in humans. We introduce new experimental results that have interesting implications for SFS. Our data favor the idea that human SFS operates in at least two distinct modes. In one mode, perceived slant is linear to luminance or close to linear with some perturbation. Whether or not the linear relationship is achieved is influenced by the relative contrasts of edges bounding the luminance variation. This mode of operation is consistent with collimated lighting from an oblique angle. In the other mode, recovered surface height is indicative of a surface under lighting that is either diffuse or collimated and frontal. Shape estimates under this mode are partially accounted for by the "dark-is-deep" rule (height $\propto$ luminance). Switching between these two modes appears to be driven by the sign of the edges at the boundaries of the stimulus. Linear shading was active when the boundary edges had the same contrast polarity. Dark-is-deep was active when the boundary edges had opposite contrast polarity. When both same-sign and opposite-sign edges were present, observers preferred linear shading but could adopt a combination of the two computational modes.

Introduction

Gradual or smooth luminance variations can give rise to the appearance of 3D undulations, even when they are not generated by realistic surface models (Aks & Enns, 1992; Kingdom, 2003; Kleffner & Ramachandran, 1992; Pentland, 1989; Schofield, Hesse, Rock, & Georgeson, 2006; Schofield, Rock, Sun, Jiang, & Georgeson, 2010; Tyler, 1998). This visual function is normally accredited to our ability to infer 3D structures from shading (shape from shading or SFS).

The precise computational mechanism underlying human SFS is poorly understood. Some early SFS algorithms in computer vision (Horn, 1975; Ikeuchi & Horn, 1981; Pentland, 1984) were also taken as candidate computational theories for human vision but were proved invalid by later experiments (Johnston & Passmore, 1994a; Mamassian & Kersten, 1996; Mingolla & Todd, 1986). A notable distinction is that classic SFS computer vision algorithms require knowledge of the light source direction (both tilt and slant angles) whereas in humans light source estimation and shape perception appear to be somewhat independent processes (Mamassian & Kersten, 1996; Mingolla & Todd, 1986; but see also Morgenstern, Murray, & Harris, 2011 who interpret the independence as a consequence of two light estimation processes: explicit and implicit, based on human data drawn from a study estimating *only the tilt angle* of light source directions). More support for this independence can also be found in studies of global shading where the addition of information on light source direction provided by global shading did not improve the accuracy of local shape judgments (Erens, Kappers, & Koenderink, 1993), although it can help resolve the concave/convex ambiguity (Berbaum, Bever, & Chung, 1983; Koenderink & van Doorn, 2004). This independence is interesting because the information on surface orientation conveyed by shading is an angle relative to the light source (Horn, 1975, 1977; Horn & Brooks, 1989)—surface orientation cannot be uniquely determined without a reference light source, yet SFS functions in humans even without such a reference.

Koenderink and van Doorn (1980) proposed that shape might be derived from an illumination and viewing direction invariant property of the distribution of image intensities. Unfortunately, humans do not demonstrate shape constancy under changing illumination (Christou & Koenderink, 1997; Khang, Koenderink, & Kappers, 2007; Koenderink, van Doorn, Christou, & Lappin, 1996a, 1996b; Nefs, Koenderink, & Kappers, 2005, 2006; Todd, Koenderink, van Doorn, & Kappers, 1996), so cannot base their shape judgments entirely on an illumination invariant property of the image.

A commonly held (but often only implicitly articulated) view in the study of SFS is that perceived slant is linear to the luminance values in the shading pattern (slant ∝ luminance). To our best knowledge, the validity and robustness of this relationship have not been tested, although it is related to the linearity theory proposed by Pentland (1989). According to Pentland, when lighting is oblique, the Fourier spectrum of surface height is linearly related to the Fourier spectrum of image intensity and their phase terms differ by 90°. Testing human observers with images of sinusoidal gratings showed that the fundamental frequency of the perceived surface equaled the frequency of the luminance grating, consistent with the theory. However, the results only provided partial proof because the 90° phase shift was not verified in Pentland's experiment. If Pentland's theory were true, then the surface perceived when viewing a sinusoidal grating ought to be a sinusoid corrugation with the same frequency as the grating but with a 90° phase shift. This phase shift will place points of maximum absolute slant at the luminance maxima and minima. If we assume positive slant to lie in the direction of the perceived illuminant, then slant is proportional to luminance. Schofield et al. (2006) mapped perceived surface profiles in response to sinusoidal shading patterns. Perceived surfaces were approximately sinusoidal corrugations at the predicted frequency, but the phase shift was typically only 45°, less than that predicted by Pentland's theory.

In some circumstances, perceived surface height (not slant) is linear to luminance—the "dark-is-deep" rule. Evidence in favor of this rule has been found in a range of shape-from-shading tasks but is most pronounced when the light is close to the viewing direction (Christou & Koenderink, 1997) or diffuse (Langer & Bülthoff, 2000). To some extent, the dark-is-deep rule is descriptive of a shading model under diffuse lighting (Langer & Zurcker, 1992). According to this model, image intensities generated under diffuse lighting depend on how much a surface position is exposed to the "sky." Thus, a sinusoidally corrugated surface will generate a luminance trace that is a periodic grating with the same fundamental frequency and phase as the surface (Wright & Ledgeway, 2004; see also Figure 1a). However, in some other cases, dark-is-deep does not fully capture shading under a diffuse illuminant. For example, in the case of a single cycle of a sine wave (Figure 1b), although the top half of

the surface obeys dark-is-deep, the bottom half of the surface has a near uniform luminance profile. Figures 1c and 1d show two further examples where the diffuse lighting model does not follow the dark-is-deep rule. Further, human SFS under diffuse light does not faithfully follow the dark-is-deep rule (Langer & Bülthoff, 2000). Langer and Bülthoff showed that small positive variations in luminance at surface troughs were not seen as localized humps as would be the case under dark-is-deep; rather they are seen as troughs. It is as if human vision knows that such deviations from dark-is-deep are possible even under diffuse illumination and hence is able to discount them.

Langer and Bülthoff (2000) showed that when surface rendering switches from collimated, oblique lighting to diffuse lighting people switch their mode for computing shape from shading accordingly. Under collimated, oblique lighting, surface points with the highest luminance are seen as facing the light source and are thus not seen as surface peaks. Under diffuse lighting, luminance peaks and perceived surface peaks coincide. As their stimuli did not include any direct cues to the directionality of the light source, people must have used cues within the shading patterns themselves to infer the most appropriate mode for SFS. However, Langer and Bülthoff did not consider what cues might drive this mode switching behavior. Our results help us infer one of the possible cues.

Human SFS is also known to produce large individual differences with only qualitative agreement between observers (Koenderink, van Doorn, & Kappers, 1992; Todd et al., 1996). These findings suggest that attempts to establish a quantitative theory for SFS are futile. However, Battu, Kappers, and Koenderink (2007) and Koenderink, van Doorn, and Kappers (2001) have shown that despite inconsistencies between observers, participants nonetheless produced consistent shape estimates up to an affine transformation. That is, the 3D representations perceived by observers differ only by scaling and shearing. Mathematically, this is described as follows: suppose that $z_1(x, y)$ and $z_2(x, y)$ are depth functions estimated by two observers, where x and y are coordinates in the image plane, the relationship between $z_1(x, y)$ and $z_2(x, y)$ can be described by a multiple linear regression: $z_1(x, y) = az_2(x, y) + bx + cy + d$ where the constant a represents a scaling factor and b, c, and d control shearing transforms of the 3D surface. Therefore, we argue that human SFS can be expressed by

$$\hat{z}(x, y) = az(x, y) + bx + cy + d, \qquad (1)$$

where $\hat{z}(x, y)$ is the 3D shape as reported by the observer, and $z(x, y)$ is the provisional 3D surface computed from shading alone and is (we presume) a function of low-level processes and therefore more or less common to all observers. In other words, the 3D information computed from shading alone is not enough to recover a full representation of the 3D shape. We call this 3D information
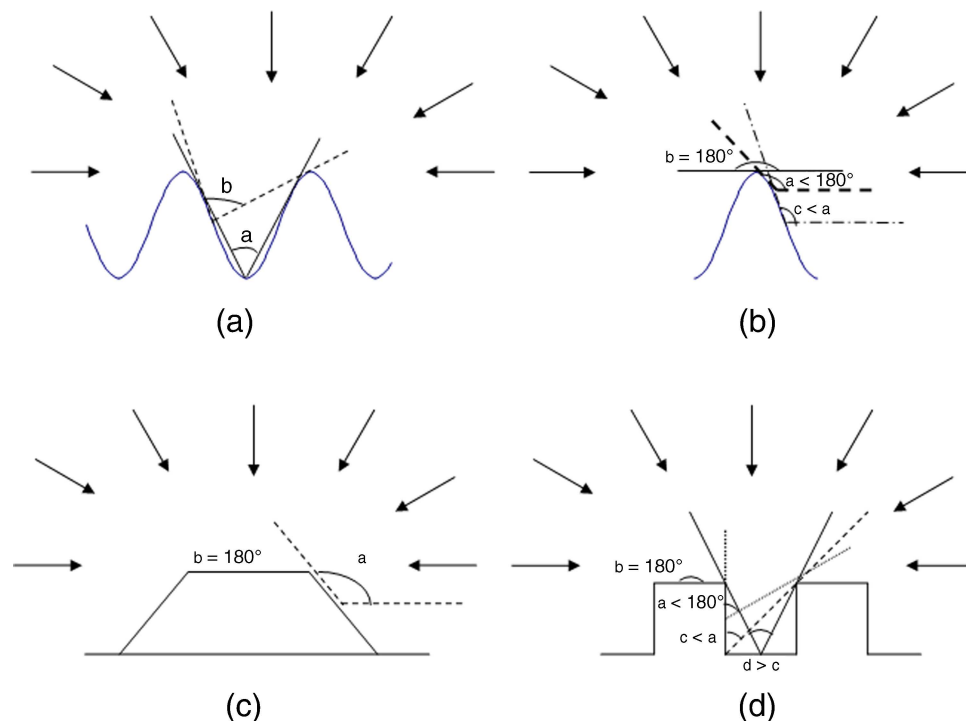
Figure 1. Validation of the dark-is-deep rule under diffuse illumination. (a) Periodic sinusoidal surface is illuminated by diffuse light. The valley sees a portion of the sky that subtends angle *a*. From the valley to the hill, the subtended angle increases and reaches the maximum at the peak, following the dark-is-deep rule. (b) A single sinusoidal bump is illuminated by diffuse light. The top of the hill sees all of the sky and, hence, is the brightest. Moving away from the peak, the surface now sees only a portion of the sky (angle *a*). This angle decreases to a minimum halfway down the bump, thereafter a small increase is observed. (c) A trapezoidal surface is illuminated by diffuse light. The top plane is exposed to the entire hemisphere while the side surface only sees part of the light source; the sides are darker than the top, but there is no change in luminance with height as one moves down the slopes. (d) A square-wave surface under diffuse light source. The top plane is exposed to the entire sky. The exposure decreases as the height of the position until the height reaches the bottom. As the measuring position moves across the valley, and if the valley is sufficiently broad, exposure to the sky increases, producing a local maximum at the center of the valley.

the "proto-surface." The remaining information is normally provided by other cues. When no other cues are available, participants "make up" for the missing information by applying their "beholder's share" (Koenderink et al., 2001), resulting in large interobserver variances. Following this logic, we believe that the common proto-surface $z(x, y)$ in Equation 1 is a key to understanding human SFS.

How might the common proto-surface be identified? In this study, we make use of the interobserver variances that are commonly found in human SFS studies. We believe that interobserver variances in SFS occur when complementary information from cues other than shading is missing or uncertain. The point at which there is not enough complementary information available to make human SFS stable (i.e., consistent across observers) is diagnostic. We hope to reveal the common processes in SFS at this point. Accordingly, in contrast to the majority of SFS studies, we chose to use stimuli made of simple luminance variations that were not meant to represent realistic 3D objects. That is, unlike most studies in human SFS, our stimuli were not generated by rendering realistic

3D objects with a predefined shading model. In doing so, we minimized the effects from other cues such as outlines (Bülthoff & Mallot, 1988; Ramachandran, 1988) and high-level object recognition.

General methods

Equipment and calibration

Stimuli were generated using VSG2/5 graphics card (Cambridge Research System, CRS, UK) and presented on a 21″ Sony Flexscan GDM-F520 CRT monitor. Responses were made via a CRS-CB3 response box connected to the VSG. Images were squares with side length of 13.3 deg (512 pixels) and were displayed inside a central window. The display was set to mean luminance outside of this window. The luminance non-linearity of the monitor was corrected using the four-parameter model proposed by Brainard, Pelli, and Robson (2002) with parameters estimated from luminance values obtained

with a CRS Colour Cal device. The viewing distance was 1 m, and the experimental monitor was the only significant light source in the room.

Stimuli and task

Most studies on human SFS use rendered or photographed objects (Battu et al., 2007; Christou & Koenderink, 1997; Khang et al., 2007; Koenderink et al., 1996a, 1996b, 1992, 2001; Langer & Bülthoff, 2000; Mamassian & Kersten, 1996; Nefs et al., 2005, 2006; Todd et al., 1996). Using realistic objects ensures observers to have a good impression of 3D shape and, hence, a fully functioning SFS system (Koenderink et al., 1996b). The logic behind this argument is that shading only conveys limited information on 3D structure (Pentland, 1984) and needs other visual information for humans to successfully articulate the underlying 3D shape (Koenderink et al., 1996b). While useful for measuring qualitative properties of human SFS, this methodology does not suit our purpose of deriving a quantitative measurement of depth computation based on shading only, as other cues such as object outlines will intervene with or dominate the shading cue (Knill, 1992; Koenderink et al., 1996a, 1996b; Mamassian & Kersten, 1996; Ramachandran, 1988).

We chose instead to use simple luminance variations that were not meant to represent realistic 3D objects. It is possible that some of our stimuli do not evoke a full 3D interpretation due to the lack of supporting information from other cues. SFS may appear to be a "broken system" for these stimuli. However, we turn this to our advantage. By testing SFS on the cusp of this broken state, we hope to titrate the contribution from shading alone. The stimuli included sinusoidal gratings, square-wave gratings, and sawtooth gratings made of repeated linear luminance ramps (see Figure 2) and variants of these waveforms where fewer cycles were presented (see Figures 2 and 5). These stimuli have been shown to produce 3D percepts in some earlier studies (Aks & Enns, 1992; Bergstrom, Gustafsson, & Putaansuu, 1984; Kingdom, 2003; Kleffner & Ramachandran, 1992; Pentland, 1989; Schofield et al., 2006, 2010). We painted the image with fine-scale textures so as to give better articulation to the 3D percepts (Sakai, Narushima, & Aoki, 2006), but the textures themselves did not provide any cues to depth—they contained no geometric distortions. We also ran a pilot experiment with and without these textures and found that they did not have any impact on participants' judgments even though participants reported that stimuli with textures looked more like real surfaces than those without.

We used a standard gauge figure task to probe the perceived surface slant (Koenderink et al., 1992). The gauge figure (shown in Figure 1a) is a 2D representation of a circular disk with a protruding stick. By varying the aspect ratio of the 2D projection of the disk and the length and direction of the stick, it is possible to create the
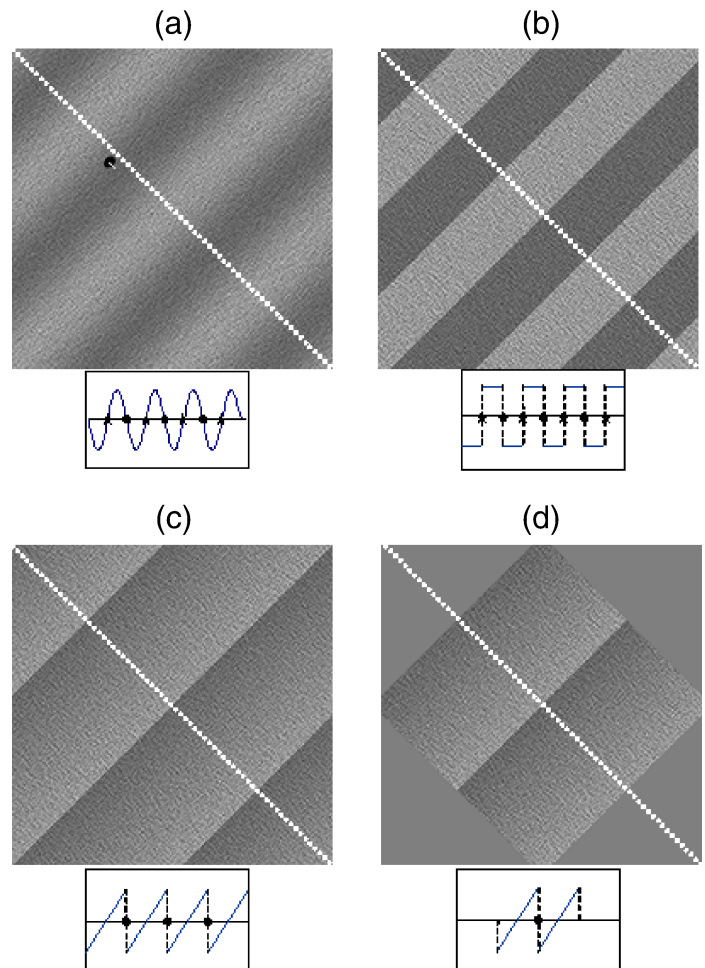

Figure 2. Sample textured luminance profiles from (a–c) Experiments 1 and (d) 2. The diagonal cross sections (white dotted lines) of their LM component are plotted below each stimulus. (a) Sine wave. (b) Square wave. (c) Sawtooth. (d) Truncated sawtooth. Stars and circles show positive and negative going edges, respectively. Edges represented by the same tokens have the same contrast and polarity. The gauge figure sued to probe surface slant is shown in (a).

impression of a slanted disk and have observers set this slant to match the perceived slant at locations in the underlying test image. Surface slant can then be integrated to derive perceived shape. The diameter of the gauge figure was 0.533 deg.

Experiment 1: Verifying slant is proportional to luminance

We first test the assumption that slant is proportional to luminance and assess the reliability of this relationship. The 3D appearance of sinusoidal, square-wave, and periodic sawtooth gratings was probed. The stimuli are

shown in Figures 2a–2c with their corresponding luminance cross section as measured on the diagonal indicated by the dashed lines. All gratings had the same minimum and maximum luminance values. All waveforms had a frequency of 0.2 c/deg and displays contained 3–4 cycles of each waveform. Note however that the sawtooth ramps are dominated by their upward slopes rather than the sharp transitions at the end of each rise. Thus these waveforms appear to have a lower frequency than the sine- and square-wave gratings. In this configuration, each stimulus contained *perceptual edges* made of either step changes in luminance or peaks in gradient magnitude, i.e., zero crossings of the second derivative of luminance (Georgeson, May, Freeman, & Hess, 2007; Hesse & Georgeson, 2005). Thus, perceptual edges are located in wherever mean luminance is achieved in sine waves but not so for linear ramps. Further, each grating in Experiment 1 had at least two edges that were equal in magnitude and contrast polarity (edge locations and their contrast polarity are indicated on the cross sections of Figure 2). Stimuli were presented at three orientations (horizontal and ±45°).

Procedure

Perceived shape was estimated from gauge figure settings. The slope of the gauge figure was randomly initialized and was adjustable only in the direction of the luminance variation. One cycle of the sine- and square-wave modulations but two consecutive cycles of the sawtooth gratings were measured. Thus, we captured perceived shape on either side of the sudden transition in the sawtooth stimuli. Probe points close to the edges in sawtooth stimuli were moved by 1/24th of a wavelength to avoid testing directly at the transition. Otherwise, the waveforms were sampled at multiples of 1/8th of a cycle of the grating (0.625 deg) with probe points randomly displaced along the orthogonal direction. Participants saw only one stimulus type in each test session, but stimuli were redrawn for each trial using a new random noise sample.

Two naïve participants and one of the authors took part in this experiment. Each made 4 settings for each test position and the mean value of the 4 gradients was taken as the perceived slant at that location. The mean gradients were integrated to estimate perceived depth in response to each stimulus.
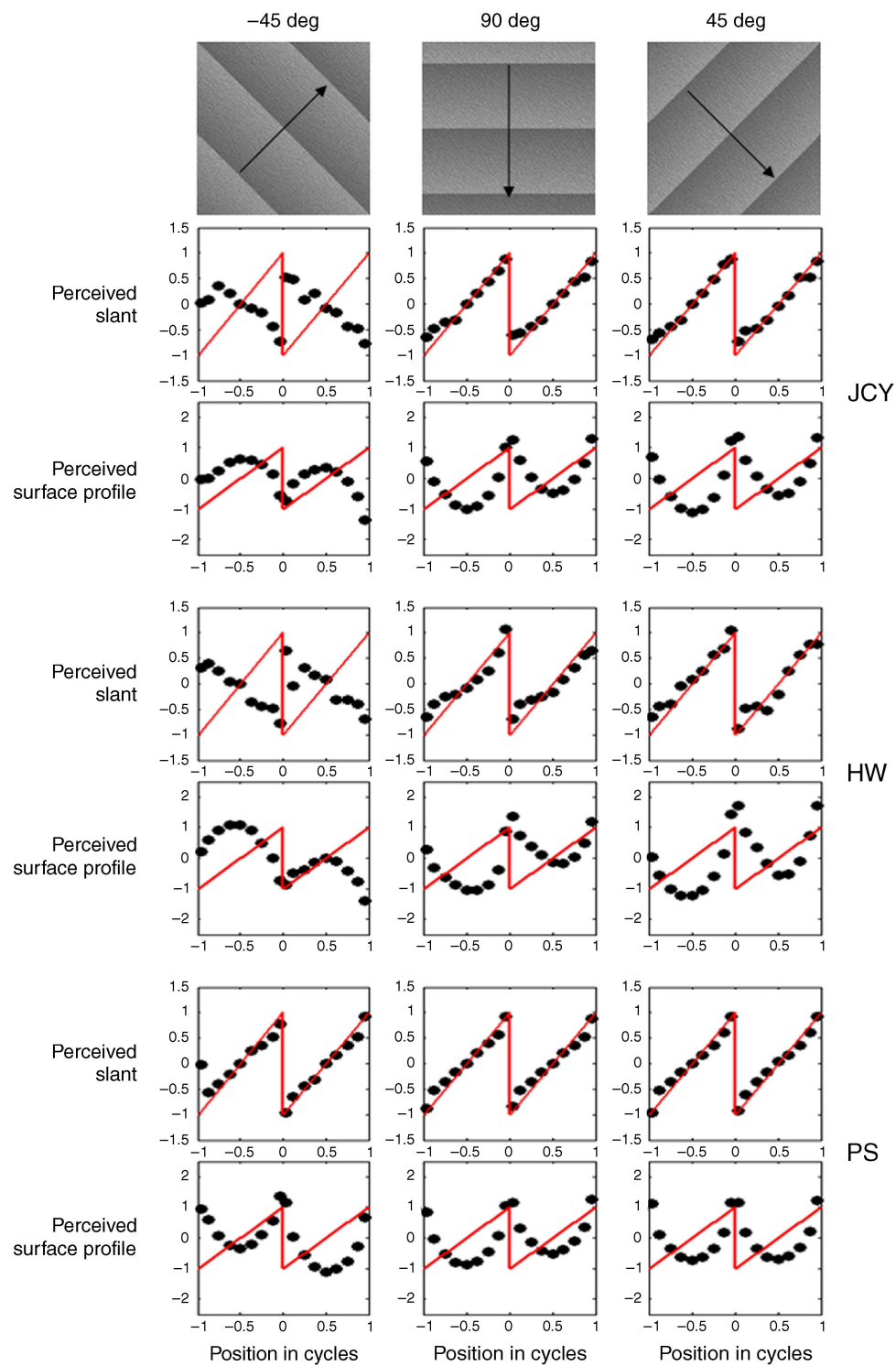
Results

Perceived slant and recovered surface profiles for all three participants are shown in Figure 3. The linear relationship between perceived slant and luminance as well as that between recovered surface heights and luminance were measured using Pearson's correlation (see Tables 1 and 2).

All three participants agreed qualitatively on the surface shape for periodical sawtooth gratings except for an ambiguity between concave and convex interpretations. The perceived slants appeared proportional to the luminance profiles of the stimuli (mean correlation = 0.96). For the two naïve observers, the sign of the relationship varied with orientation being positive for 90° and 45° but negative for −45°. These observers saw 90° and 45° sawtooth gratings as broad deep valleys with sharp ridges while the −45° sawtooth was perceived as broad mounds with sharp valleys (Figure 3a). The third participant (author PS) saw all orientations as broad valleys.

For sine-wave gratings, the two naïve subjects assumed an approximately linear relationship between perceived slant and luminance (mean correlation = 0.94). The recovered depth profiles for these two participants look like phase-shifted sine waves with a 1/4 cycle (90 degree) offset between luminance and perceived surface height. Participant PS perceived surfaces that were broadly sinusoidal but with either larger (at −45°) or smaller (45 and 90°) offsets such that at −45° the correlations between perceived height and luminance (−0.95) was stronger than that between perceived slant and luminance (0.2). At 45 and 90°, PS produced intermediate offsets between surface and luminance peaks such that it is not possible to distinguish between slant ∝ luminance and height ∝ luminance for this participant (correlations for PS's perceived slant and luminance are 0.58 and 0.67 while those for his perceived height and luminance are 0.62 and 0.5).

PS and JCY perceived 90° and 45° square-wave stimuli as triangular profile surfaces, suggesting a linear relationship between the perceived slants and luminance (mean correlation = 0.83). HW's slant estimates show no clear trends but do not suggest a linear relationship between perceived slants and luminance for 90° and 45° square-wave stimuli (mean correlation = 0.25). However, HW's relationship between the recovered surface height and luminance is roughly linear (correlation = −0.73) although the percept is clearly weak. The recovered surface heights for −45° square wave show no clear trends for any participant: perceived slants distributed around zero.

To summarize, from Table 1, data for sawtooth gratings are consistent across all participants. All three people set perceived slant proportional to the luminance profile. The sign of the relationship varied with the orientation of the grating and across participants. Sine-wave gratings were also perceived as depthy, with the perceived slant roughly proportional to the luminance profile under most conditions. Participant PS is an exception for his perceived height rather than slant correlated with luminance for the −45° grating. For 90° and 45° gratings, PS perceived a sinusoidal surface with a phase shift that was less than 90°. The square wave appeared the least corrugated of the three stimulus types. Responses for the −45° square wave suggest that no clear depth percept was obtained for this stimulus.
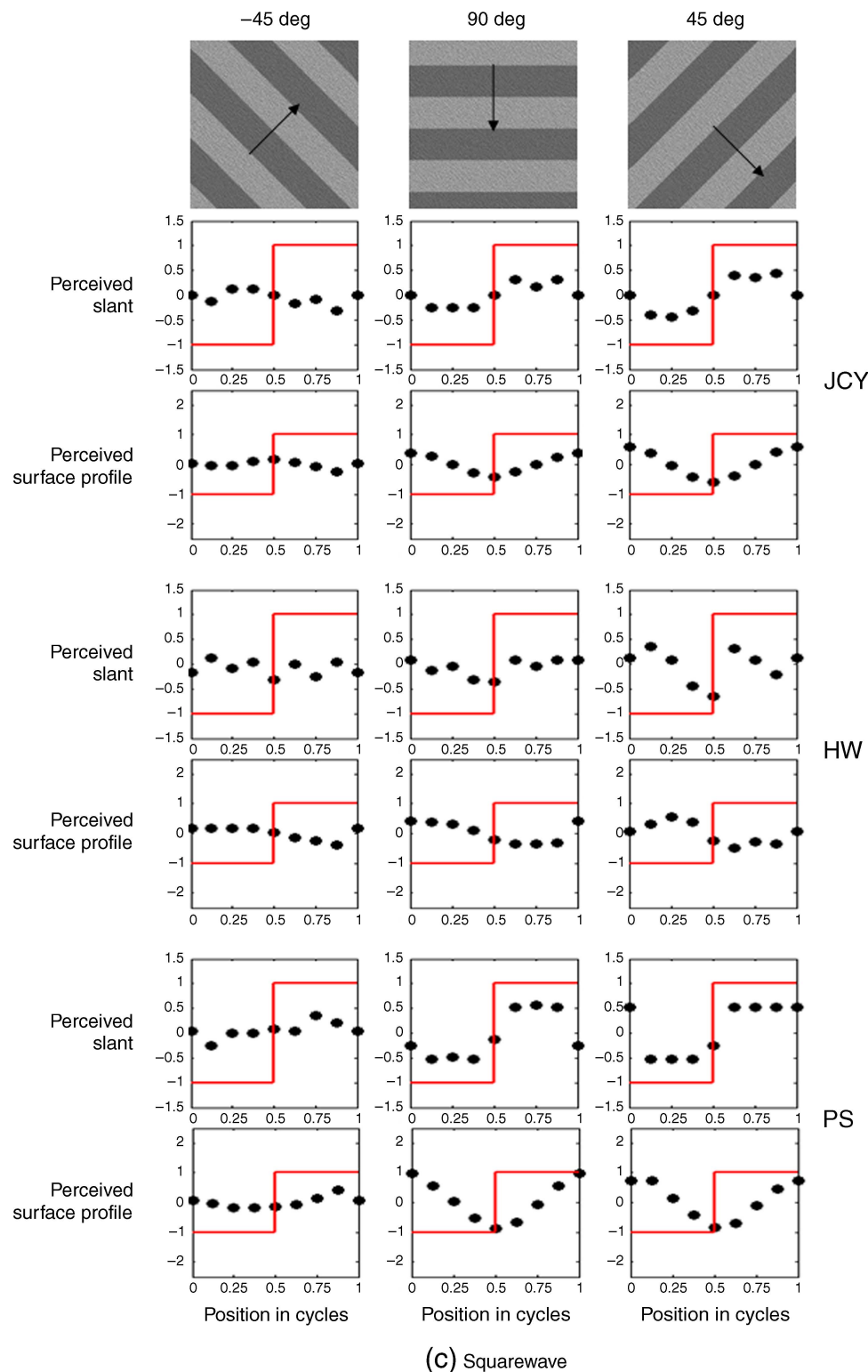
(a) Sawtooth
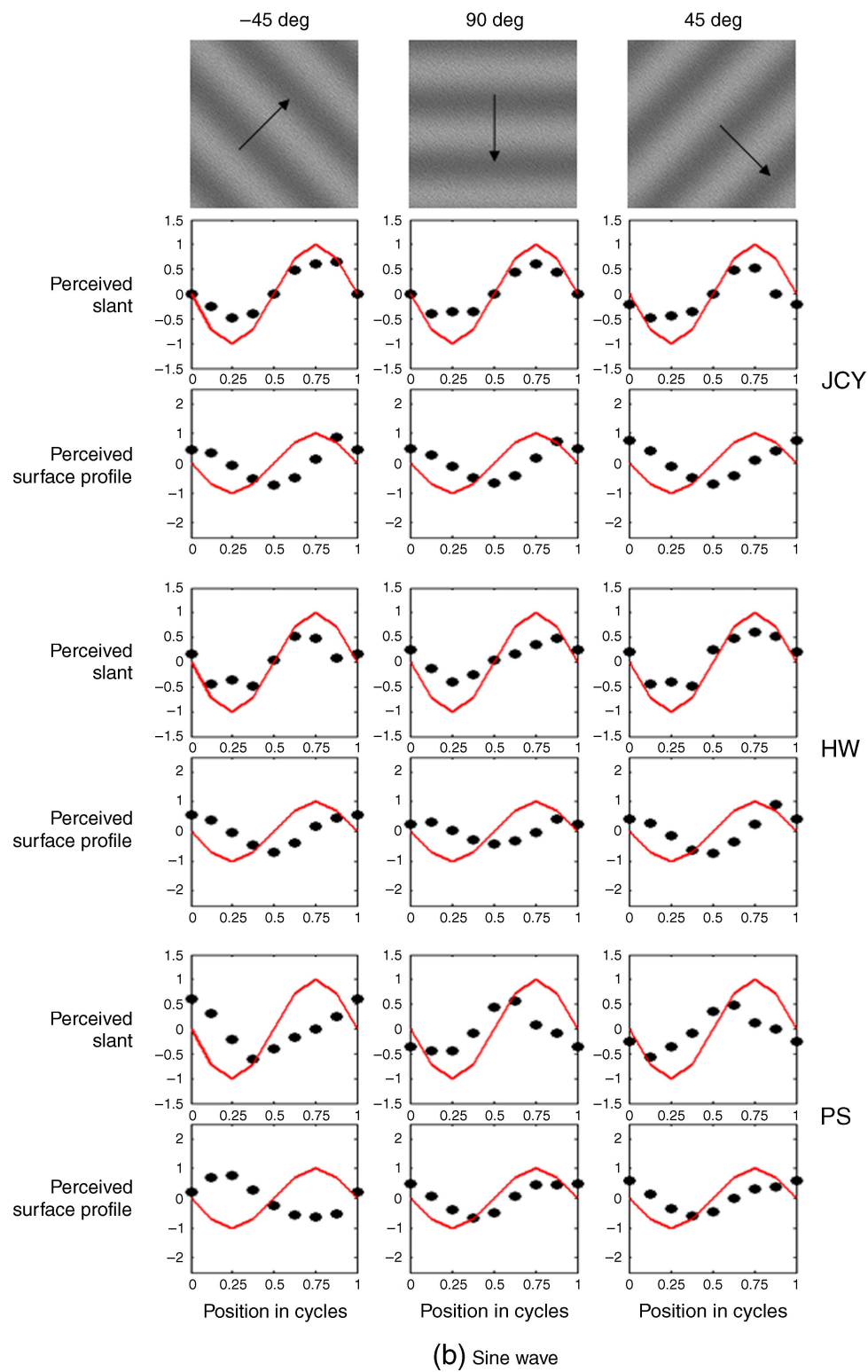
(b) Sine wave

Figure 3. Three participant's perceived slant and perceived surface profile for (a) periodical sawtooth, (b) sine-wave gratings, and (c) square-wave gratings. Results for stimuli with the same orientation are grouped in the same column. Solid, red lines represent the luminance profile. The observer's response is represented by dots. The horizontal axis is the spatial location in the unit of grating cycles. The black arrow indicates the direction of the luminance variation.

	Sine wave			Sawtooth			Square wave		
Participants	−45°	90°	45°	−45°	90°	45°	−45°	90°	45°
JCY	0.98	0.99	0.90	−0.9	0.99	0.99	−0.61	0.85	0.86
HW	0.9	0.9	0.96	−0.94	0.97	0.97	−0.25	0.44	0.06
PS	0.2	0.58	0.67	0.93	0.99	0.99	0.66	0.84	0.75

Table 1. Correlation between perceived slant and luminance.

Discussion

The results of Experiment 1 suggest that SFS is reasonably well modeled by the "slant is linear to luminance" or linear shading rule (cf. Pentland, 1989) in our reduced stimuli. The sign of this relationship varies between observers and across orientations. There are some cases/individuals for which this rule holds less strongly; here, observers seem to set surface height proportional to luminance but note that they do not follow the dark-is-deep rule as in these cases the correlation between height and luminance is negative (dark-is-deep would produce a positive relationship).

We think that variations in the relationship between slant and luminance across orientations and observers might be caused by the location of their assumed light source. In evaluating SFS, most people assume an implicit light source that is above their heads and a little to the left, but individual differences can be quite large and some people assume above right lighting (Mamassian & Goutcher, 2001). If slant is proportional to luminance, bright regions should be seen as slanted toward the light source, and thus, the sign of the relationship may depend on stimulus orientation for some observers.

We also note that there is a tendency for the relative depths in our stimuli to flip. This is because a valley lit from one side produces the same luminance profile as a hill lit from the other side. Such flipping should be stabilized by people's preferred lighting direction, but this is clearly plastic. Observers may express a preferred rather than an absolute interpretation. Since a new display was redrawn for each trial, observers may flip their surface interpretation for each new stimulus. This will weaken our relationships and may be the result of the ambiguous results for the square-wave stimuli. Participant HW's perceived slants in response to 90-deg square waves did indeed fall in two categories. Such bimodality is not as clearly present in other conditions/observers. In Experiment 2, we made the stimuli stay on the screen for the whole session and only redrew the gauge figure at each new test positions. In this way, we hope to obtain more reliable output for any particular luminance profile.

Another reason why square waves may have led to poor depth profiles might be that these stimuli have an abundance of sharp edges separating regions of uniform mean luminance. Such features might promote segmentation into regions of different surface material (see Kingdom, 2008 for a review) such that the square-wave stimuli are simply seen as flat. The fact that this occurs for square waves but not sine or sawtooth functions supports the idea that the latter are seen as genuinely corrugated.

The central idea of our methodology is to use consistencies and inconsistencies between observers and stimuli to expose the underlying computational structure for human SFS. Following Experiment 1, there are two possible ways forward. We can attempt to break the consistently strong relationship between slant and shading found for sawtooth gratings by removing some information from the stimuli (Experiment 2), or we can manipulate the square- and sine-wave gratings to make perceived shape more consistent across observers (Experiment 3).

Experiment 2: Effect of edges in SFS

Edges can arise from a variety of causes: (1) reflectance changes, (2) discontinuities in depth such as occluding boundaries, (3) discontinuities in surface orientation, and (4) illumination effects such as shadows and highlights (Marr, 1982). Edges caused by reflectance changes are often excluded by the visual system from contributing to

	Sine wave			Sawtooth			Square wave		
Participants	−45°	90°	45°	−45°	90°	45°	−45°	90°	45°
JCY	0.2	0.23	0.09	−0.23	0.05	0.04	−0.28	0.002	0.03
HW	0.1	0.004	0.34	−0.29	−0.03	0.17	−0.73	−0.66	−0.81
PS	−0.95	0.62	0.5	−0.05	0.07	0.03	0.58	0.04	0.16

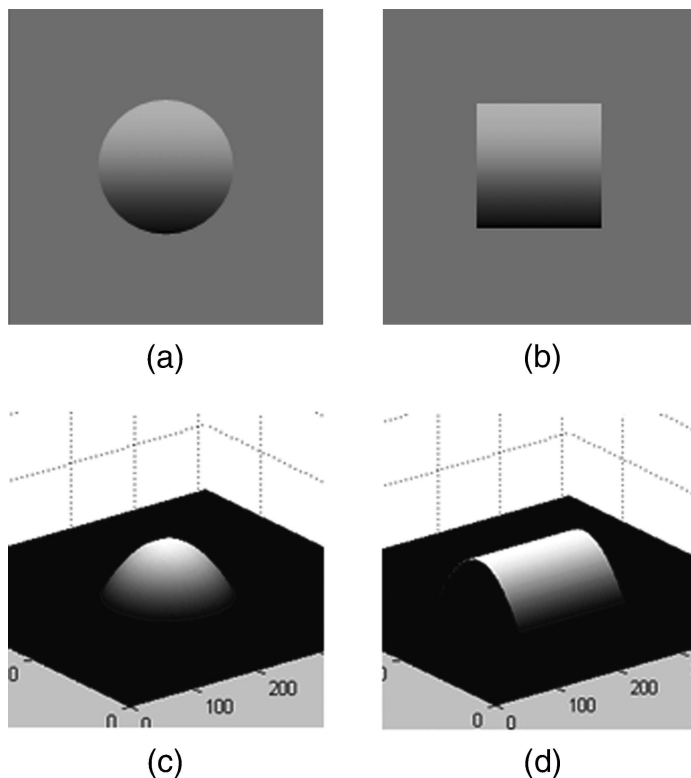Table 2. Correlation between perceived height and luminance.

Figure 4. One-dimensional luminance gradient can be perceived as (a) an ellipsoid when bounded by a circular contour but as (b) a cylinder when bounded by a square. Panels (c) and (d) were obtained by solving the ordinary differential Equation 2 with equal boundary conditions at their boundaries incepting luminance gradient. Results were produced using a simple algorithm based on the descriptions in the Discussion section.

SFS (Kingdom, 2008; Schofield et al., 2006, 2010). Occluding boundaries are a direct result of discontinuities in depth but can be a cue to surface orientation. Edges falling into this category are the points where the surface normal is perpendicular to the viewing direction (Barrow & Tenenbaum, 1981; DeCarlo, Finkelstein, & Rusinkiewicz, 2004; Malik, 1987; Marr, 1982). Edges due to changes in surface orientations are more relevant in the context of shape from shading. Edges of this type are formed by the same principle as shading gradients and can be understood as special instances of shading for which the variations in luminance are more abrupt as they arise from discontinuities in surface orientation.

Edge types 2, 3, and 4 above constitute object boundaries and edge contours that together we term outlines. Object outlines are important cues to surface shape (Ramachandran, 1988; Todd, 2004) and can be exploited to compute the 3D shape of an object (Barrow & Tenenbaum, 1981; Clows, 1971; Guzman, 1969; Malik, 1987; Marr, 1982; Waltz, 1975). The shape cue provided by outlines is so strong that it can override other cues such as shading (Bülthoff & Mallot, 1988; Knill, 1992;

Ramachandran, 1988). In such cases, object outlines alone can produce a 3D shape percept without any shading. Indeed, humans can articulate a pictorial relief similar to that based on photographs from outlines alone (Koenderink et al., 1996a). Thus, when outlines dominate shape perception, shading appears almost immaterial and its effect is either hard to measure or completely confounded by the outlines (Mamassian & Kersten, 1996). Outlines are good candidates for the complementary visual information required to make SFS function well. The interaction between edges and shading has also been utilized in computer vision. For example, classical computational approaches for shape from shading often involve solving partial differential equations. For these methods, edges and occluding boundaries can serve as initial boundary conditions because the orientations of the surface normals at these locations are known to be perpendicular to the viewing direction (Ikeuchi & Horn, 1981). However, the complementary relationship between edges and shading has not been thoroughly examined in terms of human perception. Figures 4a and 4b illustrate the importance of outlines in the perception of SFS even when outlines alone do not support unambiguous 3D perception.

Stimuli and procedure

In this experiment, we tested the effect of edges on the perception of the sawtooth stimuli. The original sawtooth stimuli were cropped so that only two cycles were visible (Figure 2d). Compared to the original sawtooth stimuli (Figure 2c) in which each modulation cycle was bounded by step edges of the same contrast and contrast polarity, the cropped sawtooth contained only one edge inside the surface region located between the modulation cycles. The other two edges shared between the figure and the background can either be thought irrelevant to the surface (being either reflectance changes between the surface and background, occlusion edges between surface and background, or both) or be regarded as edges belonging to the surface but with different contrast from that of the central edge. The procedure was the same as Experiment 1 and the same three participants were used.

Results

Perceived slants and recovered surface profiles for all three participants are shown in Figure 5 in the same format as in Experiment 1. Pearson's correlations for the relationships between slant and luminance and height and luminance are shown in Table 3. In cases where participants previously perceived the surfaces as deep valleys (90° and 45° for JCY and HW and all three orientations for PS), gratings were now perceived as a single crease between two gently curving surfaces but not valleys as such. Gradients are proportional to luminance
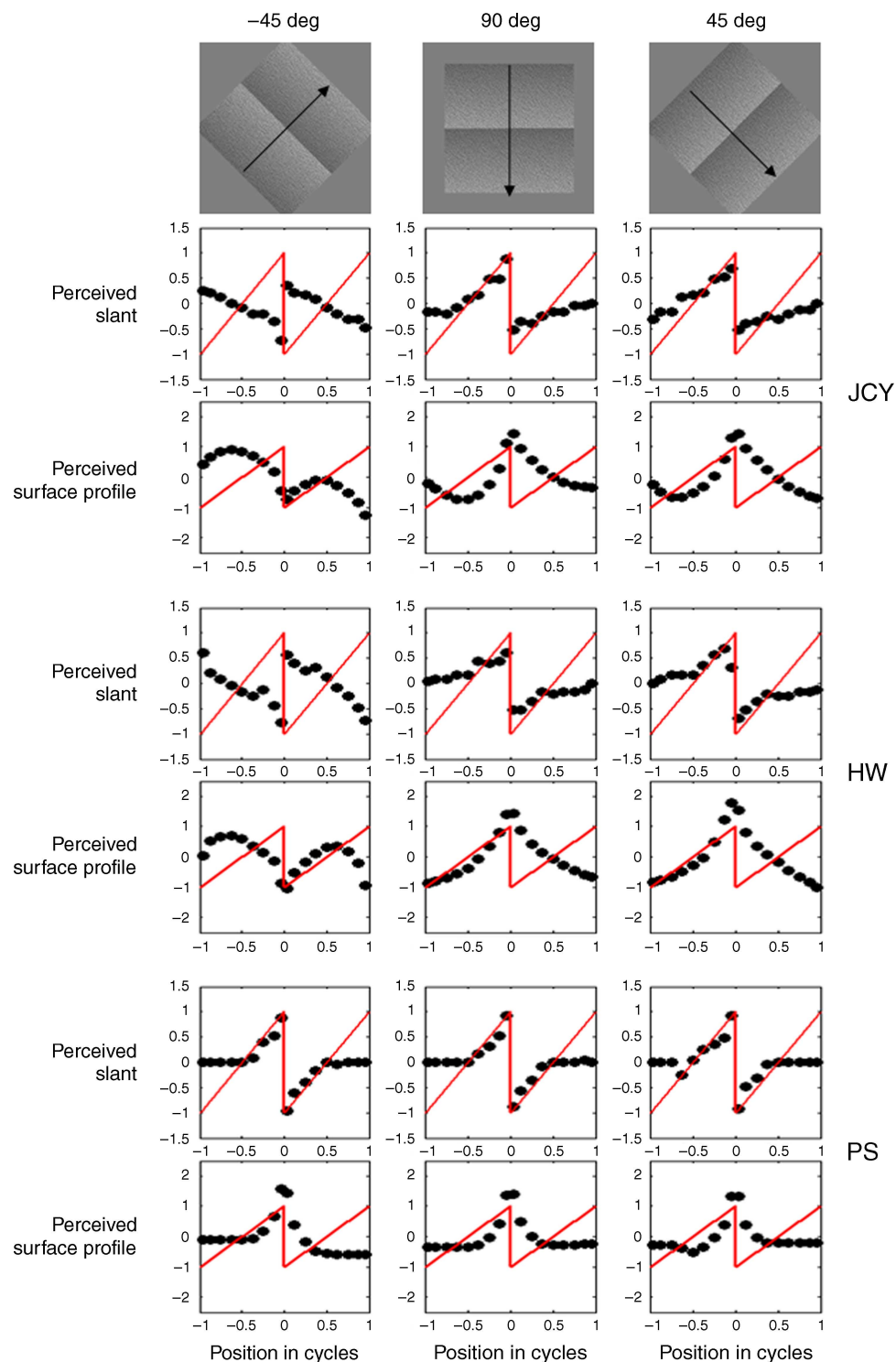
Figure 5. Shape perceptions for cropped sawtooth stimuli in Experiment 2. Details as Figure 3.

close to the central ridge but deviate from linearity toward stimulus borders. However, where observers had previously perceived broad mounds punctuated by sharp valleys (−45° grating for JCY and HW) altering the border conditions did not change the percept; gradients were still negatively proportional to the luminance (Table 3).

Discussion

Edges play an important role in the perception of SFS in sawtooth gratings. When luminance gradients were bounded by equal polarity edges (as in the case of the original sawtooth), human performance can be predicted

	Slant proportional to luminance			Height proportional to luminance		
	−45°	90°	45°	−45°	90°	45°
JCY	−0.97	0.74	0.70	−0.33	−0.17	−0.15
HW	−0.95	0.53	0.46	−0.19	0.04	0.04
PS	0.7	0.71	0.7	−0.08	0.02	0.02

Table 3. Correlation coefficients for cropped sawtooth.

by the linear relationship between slant and luminance (i.e., slant ∝ luminance) and the recovered surface were very consistent for all three observers (see Table 1 and Figure 3c). We call this the Linear Shading Model (LSM). In cropped sawtooth stimuli, the boundary conditions are undetermined: The figure–ground edges may not be included in the computation of shape from shading, or even if they are, the modulations were not bounded by equal edges because the central edge has much higher contrast. Under such condition, the linear relationship breaks down for those surfaces that had been seen as broad valleys (concavities) in Experiment 1. The recovered surface profiles suggest that the failure of the linear relationship was due to the uncertainty about the relative surface height of the two outer boundaries compared to that of the central ridge. The contrasts of the outer boundaries no longer matched that of the central ridge perhaps suggesting to the observer that the three positions were not of equal height despite representing similar points on the sawtooth cycle. Note that perceived height at the two boundaries was no longer roughly equal to the central ridge and that this height varied between participants.

It may be noted that our stimuli have some similarity with those used in the Craik–O'Brien–Cornsweet (COC) illusion raising the possibility that this effect—operating well before shape from shading—explains the results of Experiment 2. However, we argue that our results are not due to changes in perceived brightness via this illusion. In the COC illusion, relatively thin, isolated ramp edges cause differences in perceived brightness that extend across large distances such that one side of the edge is seen as much lighter than the other despite the two sides being equiluminant. Our relatively extended ramps do not produce the COC illusion. In our stimuli, the space to one side of the edge appears lighter than a similar space on the other side simply because it has a higher luminance. As you move away from the edge, both the physical and perceptual differences reverse. Further, if COC were causing our shape distortions, then we should expect similar effects for all conditions. However, our depth percepts only change (between Experiments 1 and 2) when the surfaces were perceived as concave, not convex.

We now ask, what computation might underlie the LSM to make it sensitive to boundary conditions in the way described above? It is useful to mathematically derive LSM at this point. Any assumptions or constraints

required in the derivation will provide further suggestions as to what additional information is required to make LSM operate in humans. It can be shown (see Appendix A) that when the light source is collimated and oblique, and the surface slant is less than 45°, the second derivative of the surface height is approximately linearly related to the first derivative of the image intensity:

$$z''(x) \approx CI'(x), \tag{2}$$

where C is a constant and $z(x)$ is the surface height. The solution to Equation 2 is given by

$$z(x) = C \int I(x)dx + bx + f, \tag{3}$$

or if the required output is the gradient of the surface, then the solution is given by

$$z'(x) = CI(x) + b. \tag{4}$$

We can see that the LSM is actually a family of solutions to Equation 2. In Experiment 1, participants agreed qualitatively well on the perceived slant and height for periodical sawtooth; there was little disagreements on the shearing aspect of the overall perceived heights. Mathematically, this means that Equation 2 has either Neumann boundary condition (surface slants are made certain at the two boundaries) or Dirichlet boundary condition (surface heights are made certain at the two boundaries). Here, we think that Dirichlet boundary conditions are probably the one that participants used because relative depths are normally available at edges in the real world via disparity. When observers feel unsure about the height difference and the determination of the difference in height is left to chance or the result of internal biases, then the recovered surface will be subject to individual differences leading to inconsistencies in perceived shape.

In our stimuli, disparity cues are not available; therefore, observers had to use other cues to estimate relative height. Perhaps edges with similar contrast are treated as being at roughly equal height. For cropped sawtooth gratings, the interpretation of the figure–ground segmentation between the textured surface and the gray surround is ambiguous, giving no hint as to the relative height at that position. In this case, the magnitude of the shearing component of surface height (b in Equation 3) is left completely to the individuals "beholder's share" (Koenderink et al., 2001). When surfaces appeared convex, observers still resolved the ambiguity by assigning roughly the same surface height to boundary positions, resulting in a proportional relationship between perceived slant and luminance. People may be applying additional constraints based on the physics of shapes (Pizlo, 2008). For example, two mounds resting on a single central

valley is not stable; it will fall to either side. Mounds with three points of contact at the same height are stable 3D objects. In the concave case, a central ridge can rest on flanks with any height.

Observing that Equation 3 is a 1D version of the ambiguity function for human SFS defined in Equation 1, we infer that under the LSM human SFS associates the second derivative of the height function with the first derivative of the luminance variation. In this framework, behavioral responses in SFS tasks are concerned with a specific realization of Equation 3, that is, SFS must assign values to the three coefficients based on additional visual cues in the image as well as the observers' "beholder's share." Because $z''(x)$ is a good approximation of surface curvature, this idea is consistent with the claim that, with respect to SFS, the visual system codes surface curvature, not height (Johnston & Passmore, 1994b). Note that the LSM does not require precise knowledge of the *slant angle* component of the light source direction, although it assumes that the illumination tilt angle is inline with the direction of local luminance gradient (Pentland, 1982). Such ability to compute shape without precise knowledge of the light source is presumably a desirable feature as humans readily convert shading to shape without such knowledge.

The LSM makes some useful predictions regarding the perceived 3D structure of simple luminance patterns. Suppose the luminance profiles in Figures 4a and 4b were taken as the gradient of the surface, as is the case in the LSM. Assuming that all the boundaries have the same height, column by column integration (since the direction of the luminance gradient is vertical) of Figure 4a will give rise to a series of quadratic curves with domes at different heights—an ellipsoid surface (Figure 4c). By contrast, when the same linear ramp luminance profile is bounded by a square (Figure 4b), the same process will give rise to a series of quadratic curves with domes at the same height—a cylinder (Figure 4d). Thus, in this simple example, the results of applying the LSM agree with subjective experience. The computation for complex 2D shading patterns is more complicated but shares the same principle: One can first find the direction in which the first derivative of local luminance achieves a maximum (i.e., gradient direction) and then integrate the luminance gradient along that direction. Alternatively, one can solve the 2D version of Equation 2 (i.e., the Poisson's equation) with careful choice of boundary conditions using well-developed numerical methods (e.g., Blake, 1985).

Experiment 3: Effect of the sign of edge contrast (edge polarity) in SFS

In Experiment 2, we broke the consistency of shape perception for the sawtooth stimuli by altering the

boundary conditions (surface edges) and we thus identified the LSM as a candidate for the underlying mechanism in SFS. Now we turn to the inconsistencies found for the sine- and square-wave stimuli in Experiment 1 for which behaviors other than those predicted by LSM were found. In some cases, observers seemed to set surface height proportional to luminance, although this did not necessarily follow the dark-is-deep rule as the relationships could be negative. A notable distinction between sine-/square-wave stimuli and the multiple cycle sawtooth gratings in Experiment 1 is their edge distributions. Each of the stimuli in Experiment 1 contained either step edges in luminance or edges defined by zero crossings of the second derivative of luminance. All stimuli had at least two edges that were equal in magnitude and contrast polarity. Sine- and square-wave gratings have at least 2 edges of each polarity, whereas all the edges in the sawtooth gratings had the same polarity. Perhaps edge polarity, or the distribution of edge types, influences SFS, with same-polarity edges promoting the use of the LSM by human vision. In this experiment, we investigate the effect of edge polarity to see if consistent shape perception can be (re)established by manipulating the distribution of edge types and also to see if this distribution determines which computations are used (e.g., LSM vs. dark-is-deep).

Methods

The stimuli were made from the same sine- and square-wave gratings as in Experiment 1. Some gratings were cropped such that the retained section contained 1.2 cycles of modulation. Thus, the only remaining visible edges in the figure had opposite polarities (see Figures 6c and 6d). Stimulus orientation was fixed at 45°. Perceived shape was measured using the gauge figure task of Experiment 1 except that the disk had a smaller diameter of 0.48 deg. The adjustment steps for the gauge figure were made either 1° or 10° so that observers could toggle between coarse and fine adjustments. The measuring points were sampled at multiples of 1/10th of a cycle of the grating (0.5 deg) but randomly displaced along the orthogonal direction. Thus, the diameter of the disk (0.48 deg) was less than the sampling distance (0.5 deg). The measuring positions were arranged so as to avoid directly testing at edges in the square-wave gratings. Measuring positions started at 1/20th of a cycle from the top left edge of the cropped stimuli and at a similar position relative to the center of the uncropped stimuli (see Figure 6). Three new naïve participants were tested in Experiment 3.

Results

Figure 7 describes the data in a similar format to Figure 3. Table 4 gives Pearson's correlation coefficients
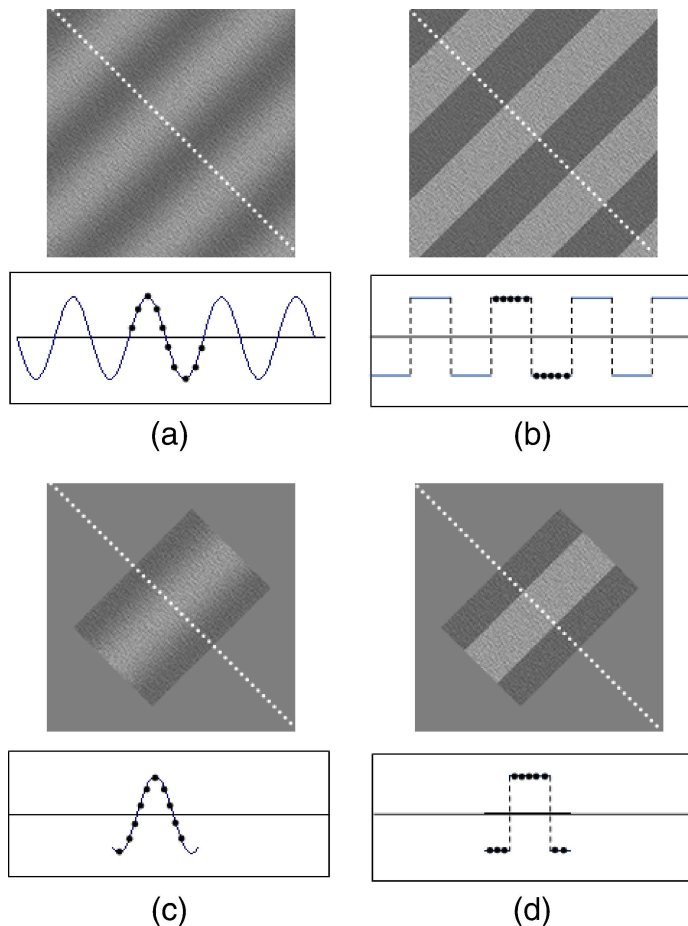
(a)   (b)

(c)   (d)

Figure 6. Stimuli in Experiment 3. (a) Sine wave and (b) square wave are the same as in Experiment 1. Panels (c) and (d) are cropped versions of (a) and (b), respectively. The visible portions in (c) and (d) are 1.2 cycles of the periodical gratings. Panels (a) and (c) and (b) and (d) are shifted by 90° in phase. The dots mark the ten measuring positions within a cycle of the test gratings.

between perceived slant and luminance and between perceived height and luminance, respectively.

Perceived surface profiles for uncropped sine-wave gratings were similar to those of Experiment 1: Two observers (TT and ZXQ) produced slant estimates linearly related to luminance (correlations = 0.98 and 0.87), whereas the correlations between their perceived heights and luminance were low (−0.44 and −0.28). The coefficients for the other observer (KL) were both moderate (0.67 and 0.58 for gradient and height vs. luminance, respectively). For cropped sine waves, no participants demonstrated a linear relationship between perceived slant and luminance (correlations = −0.23, −0.2, and −0.27). However, their correlations for perceived height and luminance were all strongly positive for the cropped sine wave (0.96, 0.76, and 0.94). When viewing uncropped square waves, all participants agreed on a linear relationship between slant and luminance (correlations = 0.98, 0.97, and 0.98). The correlations

between heights and luminance were consistently low (correlations = −0.32, −0.64, and −0.5). However, for the cropped square wave, this pattern was destroyed (correlations between gradient and luminance = −0.26, −0.23, and −0.37). Instead, perceived height and luminance correlated relatively well (0.76, 0.69, and 0.73).

Discussion

For sine-wave gratings, the LSM will produce a sinusoidal surface with a 90° phase shift to the luminance (cf. Pentland, 1989), whereas a dark-is-deep model will produce a sinusoidal profile that is in phase with luminance. The perceived shape of uncropped sine wave could be explained by the LSM for two observers; however, they both switched to a dark-is-deep interpretation when the sine wave was cropped to produce a single luminance peak. From the plots (bottom left in Figure 7a), shape judgments for participant KL also appeared as a sinusoidal surface for the periodic grating but with a smaller phase shift than predicted by the LSM, as if they were using a combination of the two model predictions (a linear combination of LSM and dark-is-deep will produce an intermediate phase offset between perceived surface peaks and luminance peaks for sine-wave stimuli). However, this participant also switched to dark-is-deep when judging cropped sine waves. In this case, the dark-is-deep interpretation is consistent with both a diffuse lighting assumption and collimated, frontal lighting.

For uncropped square waves, performance is well explained by the LSM. Cropping these stimuli such that only two, opposite polarity, edges remained made all observers change their strategy to something approaching dark-is-deep. Although correlations between luminance and perceived surface height increased significantly for cropped square waves, they were not as high as for cropped sine waves. It is also clear from Figure 7 that perceived surfaces had trapezoidal rather than square cross sections. However, a cropped square-wave luminance profile is consistent with a diffusely lit trapezoidal ridge (see Figure 1c) and with a trapezoidal ridge lit from the front. Once again, cropping the stimuli caused observers to switch to a shape recovery model that is consistent with either diffuse or frontal lighting. Note that the "height $\propto$ luminance" rule reported by Christou and Koenderink (1997) is most pronounced when the direction of the light source was close to the viewing direction (frontal lighting).

Whether the strategy for cropped square waves was the same as that for cropped sine waves is open to discussion. The perceived shapes of these two types of stimuli were qualitatively similar except that one was smoothly curved and the other was made of planar surfaces. Considering the similarities of the two luminance traces, it is possible that observers switched to the same strategy when only opposite polarity edges were present. If this were so, the
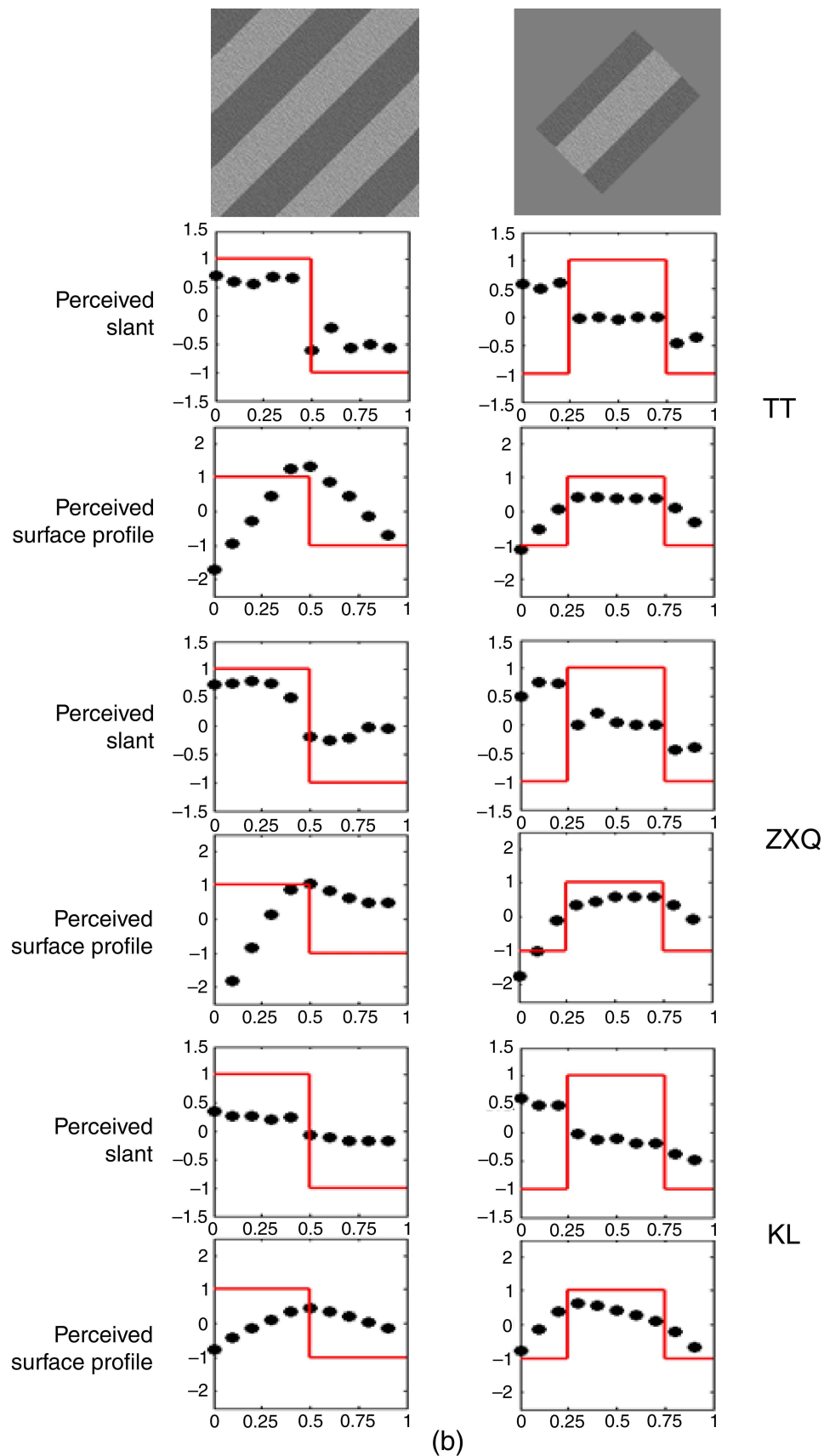
(a)
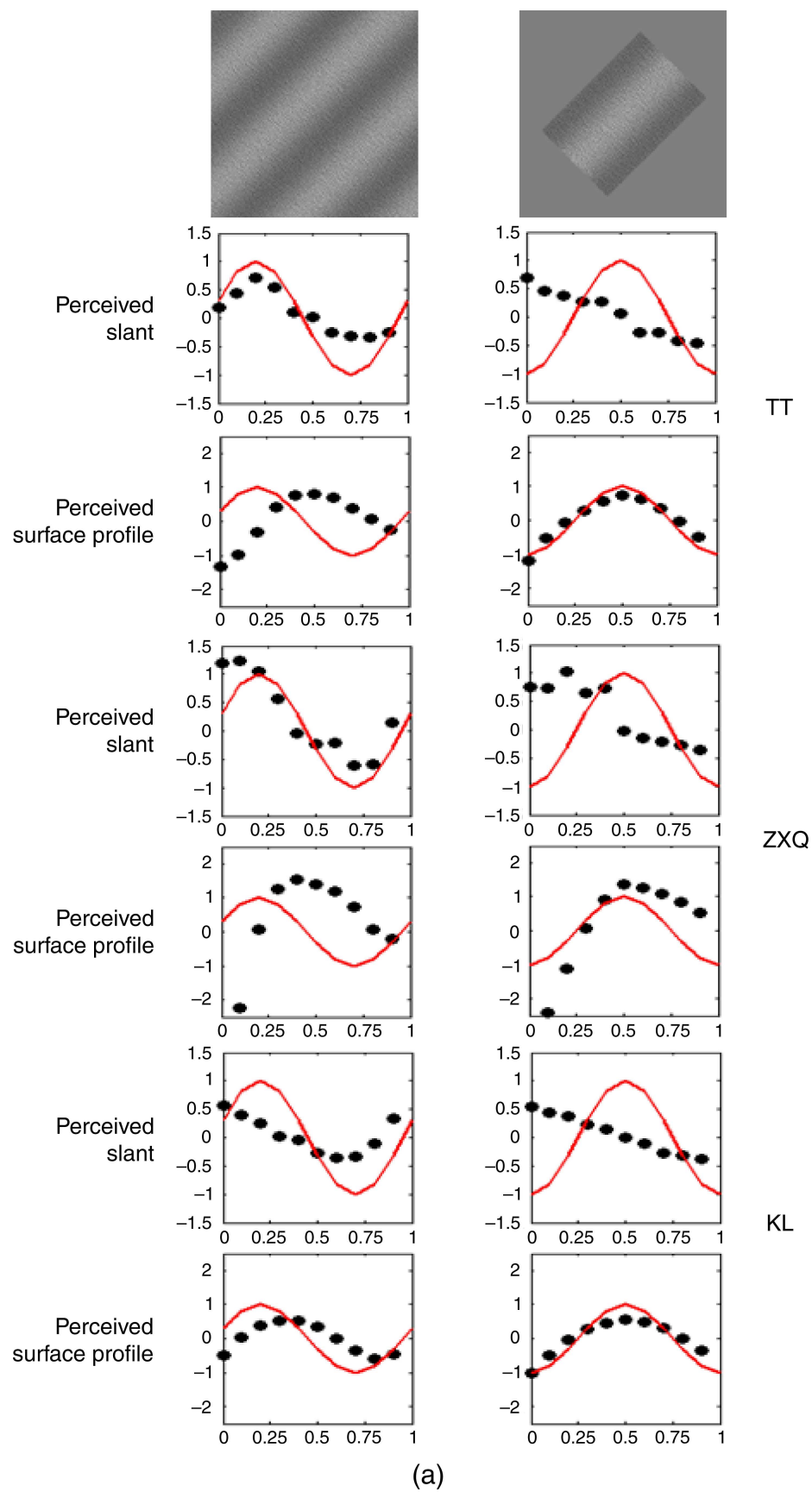
Figure 7. Three participants' perceived slants and perceived surface profiles for (a) sine-wave gratings and (b) square-wave gratings. Legends are the same as in Figure 3.

Participants	Sine		Cropped Sine		Square		Cropped square	
	Gradient	Height	Gradient	Height	Gradient	Height	Gradient	Height
TT	0.98	−0.44	−0.23	0.96	0.98	−0.32	−0.26	0.76
ZXQ	0.87	−0.28	−0.2	0.76	0.97	−0.64	−0.23	0.69
KL	0.66	0.58	−0.27	0.94	0.98	−0.5	−0.37	0.73

Table 4. Pearson coefficients between each observer's perceived gradients and the luminance, as well as between perceived surface heights and luminance for all stimuli.

dark-is-deep rule would not serve as a perfect model to characterize the unknown strategy, though it might provide an approximate model and a good fit in many cases (note that Langer & Bülthoff, 2000, found only an approximate correspondence between human shape judgments and the dark-is-deep rule in their diffuse lighting condition). What is certain, however, is that the alternative strategy matches a lighting assumption that is either diffuse or, if collimated, frontal to the image plane rather than oblique.

General discussion

To explain our data in light of other studies, we argue that human SFS operates in two modes each associated with different lighting patterns (see also Christou & Koenderink, 1997; Langer & Bülthoff, 2000; Nefs et al., 2005, 2006). When the light source is presumed to be collimated and oblique, the LSM is deployed, and in many cases, perceived slant is proportional to luminance although deviation can occur when there are insufficient constraints to solve Equation 4. When the illumination is presumed to be either diffuse or frontal to the surface, an alternative regime is used, which often (but not always) leads to perceived height being proportional to luminance (i.e., dark-is-deep). It is quite possible that there are two systems for SFS implementing the two modes described above in parallel and that the balance between these systems is determined by lighting cues in the image. Thus, it may be possible for humans to perceive surfaces that are intermediate between the two lighting interpretations such as when sinusoidal gratings are perceived as sinusoidal surfaces whose peaks are offset from the luminance peaks but by less than 90 degrees of phase offset that is predicted by the LSM. There may be many factors affecting the balance between the two mechanisms (e.g., the tilt angle of the estimated lighting direction; Schofield, Rock, & Georgeson, 2011), but in our data edge polarities seem to play an important role. Luminance variations bounded by edges with the same polarity are likely to favor the LSM, but otherwise a variant of the dark-is-deep rule might prevail. Note that once this rather coarse distinction

between oblique and diffuse/frontal lighting is made, explicit knowledge of the light source direction is not needed for either computation except to resolve the hill, valley ambiguity (in the LSM), which itself requires only a very coarse assessment of the light source direction.

Perhaps the commonly held view that slant is proportional to luminance may arise as a specific, if quite common, instance of one of the two computations outlined above: LSM. Under this mode, surface curvature is coded as the first derivative of luminance. Solution to SFS provided by this mechanisms using shading alone will not be unique; other information (e.g., relative heights of the boundary positions) is needed to further constrain the process and stabilize the perceived surface interpretations both within and between observers. We postulate, based on Experiment 2, that luminance edges provide *one* source of suitable constraints.

Note that the computations we propose here do not involve estimates of convexity/concavity since the sign of surface curvature cannot be determined given shading alone (Pentland, 1984). Therefore, the visual system still needs constraints (alternatively priors) such as overhead lighting and preference for object convexity to complete the whole process of SFS.

The LSM associates surface curvature with the first derivative of luminance and is consistent with the claim that the visual system maps shading to surface curvature rather than gradient per se (Johnston & Passmore, 1994b). The solutions given by LSM are ambiguous in the sense that the scaling and shearing factors are undetermined. However, human SFS behaves in a similarly ambiguous way (Battu et al., 2007; Koenderink et al., 2001) with perceived shape being consistent across observers only up to affine transformations.

Our results suggests that people may use a single computation (LSM) to interpret surfaces lit by a wide range of lighting scenarios as long as they perceive the light to be collimated and oblique. Appendix B shows that the LSM will overestimate the slant of a Lambertian surface when the surface is only slightly slanted but will underestimate it when the actual surface slant gets larger. The degree of underestimation will increase with increased physical slant. Similar behavior was also reported for human SFS (Hann, Erens, & Noest, 1995; Mamassian & Kersten, 1996). Moreover, close examination of those

stimuli for which such performance was reported reveals that they represented oblique lighting conditions.

Precise knowledge of the lighting direction is not needed in our framework because light source direction is not involved in the piecewise computation of local surface shape. This is consistent with the claim that SFS is independent of estimates of light source direction (Mamassian & Kersten, 1996; Mingolla & Todd, 1986) and that perceived curvature remains constant under small changes in lighting directions as long as the lighting is not frontal (Curran & Johnston, 1994). However, under our theory, the visual system has to know whether the light source is collimated and oblique so as to choose the appropriate computation. Human SFS does appear to modulate its operational mode in response to apparent changes in the illumination pattern. For example, different behaviors have been reported for different lighting patterns (mainly oblique vs. frontal and oblique vs. diffuse) during a curvature discrimination task (Curran & Johnston, 1996; Johnston & Passmore, 1994a), surface attitude judgment tasks on rendered images (Christou & Koenderink, 1997; Langer & Bulthoff, 2000; Nefs, 2008), and surface attitude judgment tasks for photographs of real objects (Todd et al., 1996).

We do not fully specify an algorithm for deciding whether or not the light is collimated and oblique. However, for simple images like those used here, the decision seems to be based on the polarities of edge pairs bounding the luminance variations in question. For natural images, it is possible that switching between the operational modes is cued by distributions of edge types. For example, there is evidence suggesting that the activities of edge detectors in response to complex images made up of Gaussian textures can be decisive in light field estimation tasks (Koenderink, van Doorn, & Pont, 2007).

Shading is inherently ambiguous. For each possible lighting direction, there exists a corresponding surface in a family of affine transformation to generate the same shading pattern (bas-relief ambiguity; Belhumeur, Kriegman, & Yuille, 1999). Thus, it is more plausible for human SFS to interpret shading in terms of a set of 3D surfaces than to achieve a unique surface representation with a specific light source direction. Given shading alone, human SFS must derive a family of functions to describe the 3D shape and then apply further constraints to choose from among the family of solutions. This strategy allows more freedom for interactions with other depth cues at later stages.

Conclusion

We have identified two constraints that humans need for conducting SFS tasks: the relative heights at surface boundaries and the general pattern of illumination. We

offer one possible interpretation for our data. That is human SFS has at least two computational modes. The common "rules" of slant being proportional to luminance and dark-is-deep are special cases of the two computational modes that we have described. LSM can explain many aspects of human SFS when a collimated and oblique light source is assumed, including the deviations from the rule that the slant is proportional to the luminance. The other mode of operation is used when the light is assumed to be diffuse or frontal and is less well defined but can be approximated by dark-is-deep in many cases. We propose that the balance between the two modes of computation depends of image content in a complex way but that edges, specifically the polarity of bounding edges, provide one cue to the composition of the illumination field.

Appendix A

Derivation of the Linear Shading Model

Assuming a Lambertian surface lit by a distant light source and viewing direction fixed to be perpendicular to the image plane, the normalized image intensity will be

$$I(x) = \cos i = \frac{\bar{n} \cdot \bar{I}}{|\bar{n}| \cdot |\bar{I}|} = \frac{p \sin \sigma + \cos \sigma}{\sqrt{p^2 + 1}}, \qquad (A1)$$

where σ is the angle between the incident ray and the viewing direction, $\bar{I} = (\cos \sigma, \sin \sigma)$ is the vector of the incident ray, p is the slope of the surface along the image plane, i.e., $p = \tan \theta$, and $\bar{n} = (1, p)$ is the vector of the surface norm (Figure A1). Note that the image plane has been simplified to be 1D in this expression. We have also omitted the multiplying constant associated with the light source. Taking the Taylor series expansion of Equation A1 about $p = 0$ up to its quadratic term will give

$$I(x) \approx \cos \sigma + p \sin \sigma - \frac{\cos \sigma}{2} p^2. \qquad (A2)$$

Pentland (1989) argued that when $|p| \ll 1$ (leading to a negligible quadratic term $\frac{\cos \sigma}{2} p^2$) and the DC term $\cos \sigma$ is ignored, the relationship between image intensity and the surface slope is linear. However, we think that omitting the DC term in Equation A2 is rather ad hoc. A more principled way to decouple the DC term from the linear term (supposing that the quadratic term $\frac{\cos \sigma}{2} p^2$ is small enough to be ignored) is to differentiate the two sides of the equation:

$$I'(x) \approx p' \sin \sigma \approx C \cdot z''(x), \qquad (A3)$$
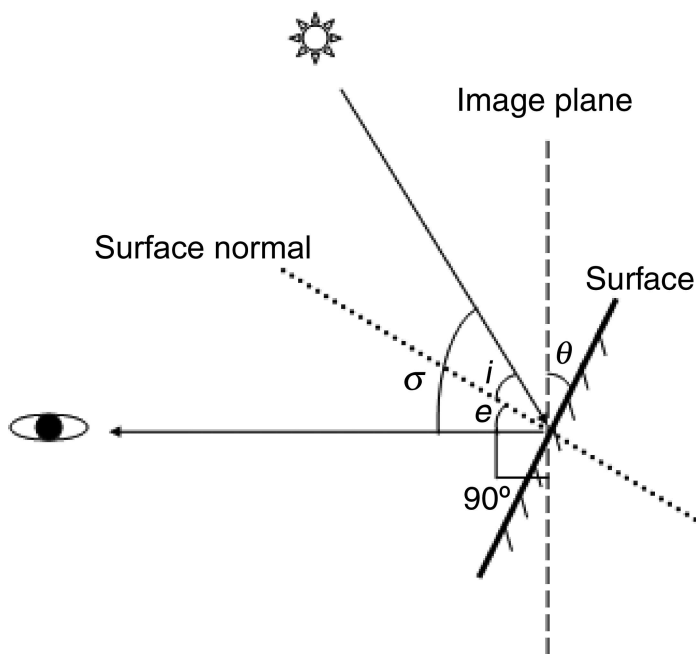
Figure A1. The relationship between image intensity (luminance) and the orientation of a Lambertian surface lit by a single point source. The process is illustrated in 1D; *e*, *i* are angles of emittance and incidence respectively. σ is the angle between the incident ray and the emittant ray; θ is the angle that the surface is inclined with respect to the image plane. Without the loss of generality, the viewing direction is set perpendicular to the image plane. Under this setting, *e* equals θ.

where C is a constant and $z(x)$ is the height function of the physical surface.

Appendix B

Systematic inaccuracies of the LSM

Within the context illustrated in Figure A1, we have $i + \theta = \sigma$. Here, we do not consider the backlit condition, so $\sigma < 90°$. Let $\hat{\theta}$ be the slant angle estimated by observers. Since perceived slant is linear to luminance, we have $\cos i = \tan \hat{\theta} \geq \sin(90° - \sigma + \theta) = \tan \hat{\theta}$. If σ equals 90°, then $\tan \hat{\theta} = \sin\theta \geq \sin \hat{\theta} = \sin \theta \cos \hat{\theta} = > \sin \hat{\theta} < \sin \theta$, so the slant angle should always be underestimated. As σ varies and let $\alpha = 90° - \sigma > 0$, then $\tan \hat{\theta} = \sin (\theta + \alpha) > \tan \theta$ when θ is very small, but $\tan \hat{\theta} = \sin (\theta + \alpha) < \tan \theta$ as θ increases and the difference becomes even larger as θ approaches 90°, i.e., perceived slant is overestimated when the Lambertian surface is only slightly slanted but becomes underestimated when the slant gets larger. The underestimation will increase with the true slant of the surface.

References

Aks, D. J., & Enns, J. T. (1992). Visual search for direction of shading is influenced by apparent depth. *Perception & Psychophysics, 52,* 63–74.

Barrow, H. G., & Tenenbaum, J. M. (1981). Interpreting line drawings as three-dimensional surfaces. *Artificial Intelligence, 17,* 75–116.

Battu, B., Kappers, A. M. L., & Koenderink, J. J. (2007). Ambiguity in pictorial depth. *Perception, 36,* 1290–1304.

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

Berbaum, K., Bever, T., & Chung, C. S. (1983). Light source position in the perception of object shape. *Perception, 12,* 411–416.

Bergstrom, S. S., Gustafsson, K. A., & Putaansuu, J. (1984). Information about three dimensional shape and direction of illumination in a square wave grating. *Perception, 13,* 129–140.

Blake, A. (1985). Boundary conditions for lightness computation in Mondrian World. *Computer Vision, Graphics, and Image Processing, 32,* 314–327.

Brainard, D. H., Pelli, D. G., & Bobson, T. (2002). Display characterization. In J. Hornak (Ed.), *Encyclopedia of imaging science and technology* (pp. 172–188). New York: Wiley.

Bulthoff, H. H., & Mallot, H. A. (1988). Integration of depth cues: Stereo and shading. *Journal of the Optical Society of America, 5,* 1749–1758.

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

Clows, M. B. (1971). On seeing things. *Artificial Intelligence, 2,* 79–116.

Curran, W., & Johnston, A. (1994). Integration of shading and texture cues: Testing the linear model. *Vision Research, 34,* 1863–1874.

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

DeCarlo, D., Finkelstein, A., & Rusinkiewicz, S. (2004). Interactive rendering of suggestive contours with temporal coherence. *Proceedings of the Third International Symposium on Non-Photorealistic Animation Rendering (NPAR) 2004,* 15–24.

Erens, R. G. F., Kappers, A. M. L., & Koenderink, J. J. (1993). Estimating local shape from shading in the presence of global shading. *Perception & Psychophysics, 54,* 334–342.

Georgeson, M. A., May, K. A., Freeman, T. C. A., & Hesse, G. S. (2007). From filters to features: Scale–space analysis of edge and blur coding in human vision. *Journal of Vision, 7*(13):7, 1–21, http://www.journalofvision.org/content/7/13/7, doi:10.1167/7.13.7. [PubMed] [Article]

Guzman, A. (1969). Decomposition of a visual scene into three-dimensional bodies. In A. Grasselli (Ed.), *Automatic interpretation and classification of images* (pp. 243–276). New York: Academic Press.

Hann, E. D., Erens, R. G. F., & Noest, N. J. (1995). Shape from shaded random surfaces. *Vision Research, 35,* 2985–3001.

Hesse, G. S., & Georgeson, M. A. (2005). Edges and bars: Where do people see features in 1-D images? *Vision Research, 45,* 507–525.

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

Horn, B. K. P. (1977). Understanding image intensities. *Artificial Intelligence, 8,* 201–231.

Horn, B. K. P., & Brooks, M. J. (1989). *Shape from shading.* Cambridge, MA: MIT Press.

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

Johnston, A., & Passmore, P. J. (1994a). Shape from shading: Surface curvature and orientation. *Perception, 23,* 169–189.

Johnston, A., & Passmore, P. J. (1994b). Independent encoding of surface orientation and surface curvature. *Vision Research, 34,* 3005–3012.

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

Kingdom, F. A. (2003). Colour brings relief to human vision. *Nature Neuroscience, 6,* 641–644.

Kingdom, F. A. (2008). Perceiving light versus material. *Vision Research, 48,* 2090–2105.

Kleffner, D., & Ramachandran, V. S. (1992). On the perception of shape from shading. *Perception & Psychophysics, 52,* 18–36.

Knill, D. C. (1992). Perception of surface contours and surface shape: From computation to psychophysics. *Journal of the Optical Society of America, 9,* 1449–1464.

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

Koenderink, J. J., & van Doorn, A. J. (2004). Shape and shading. In L. M. Chalupa & J. S. Werner (Eds.), *The visual neurosciences* (pp. 1090–1105). Cambridge, MA: The MIT Press.

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

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

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

Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. L. (2001). Ambiguity and the 'mental eye' in pictorial relief. *Perception, 30,* 431–448.

Koenderink, J. J., van Doorn, A. J., & Pont, S. C. (2007). Perception of illuminance flow in the case of anisotropic rough surfaces. *Perception & Psychophysics, 69,* 895–903.

Langer, M. S., & Bulthoff, H. H. (2000). Depth discrimination from shading under diffuse lighting. *Perception, 29,* 649–660.

Langer, M. S., & Zurcker, S. W. (1992). Qualitative shape from active shading. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR) 1992,* 713–715.

Malik, J. (1987). Interpreting line drawings of curved objects. *International Journal of Computer Vision, 1,* 73–107.

Mamassian, P., & Goutcher, R. (2001). Prior knowledge on the illumination position. *Cognition, 81,* B1–B9.

Mamassian, P., & Kersten, D. (1996). Illumination, shading and the perception of local orientation. *Vision Research, 36,* 2351–2367.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: W. H. Freeman.

Mingolla, E., & Todd, J. T. (1986). Perception of solid shape from shading. *Biological Cybernetics, 53,* 137–151.

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, 108,* 12551–12553.

Nefs, H. T. (2008). Three-dimensional object shape from shading and contour disparities. *Journal of Vision, 8*(11):11, 1–16, http://www.journalofvision. org/content/8/11/11, doi:10.1167/8.11.11. [PubMed] [Article]

Nefs, H. T., Koenderink, J. J., & Kappers, A. M. L. (2005). The influence of illumination direction on the pictorial reliefs of Lambertian surfaces. *Perception, 34,* 275–287.

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

Pentland, A. (1982). Finding the illuminant direction. *Journal of Optical Society of America, 72,* 448–455.

Pentland, A. (1984). Local shading analysis. *IEEE Transaction on Pattern Analysis and Machine Intelligence, 6,* 170–187.

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

Pizlo, Z. (2008). *3D shape: Its unique place in visual perception*. Cambridge, MA: The MIT Press.

Ramachandran, V. S. (1988). Perception of shape from shading. *Nature, 331,* 163–166.

Sakai, K., Narushima, K., & Aoki, N. (2006). Facilitation of shape-from-shading perception by random textures. *Journal of the Optical Society of America, 23,* 1805–1813.

Schofield, A. J., Hesse, G., Rock, P., & Georgeson, M. (2006). Local luminance amplitude modulates the interpretation of shape-from-shading in textured surfaces. *Vision Research, 46,* 3462–3482.

Schofield, A. J., Rock, P. B., & Georgeson, M. A. (2011). Sun and sky: Does human vision assume a mixture of point and diffuse illumination when interpreting shape-from-shading? *Vision Research, 51,* 2317–2330. [PubMed] [Article]

Schofield, A. J., Rock, P. B., Sun, P., Jiang, X., & Georgeson, M. A. (2010). What is second-order vision for? Discriminating illumination versus material changes. *Journal of Vision, 10*(9):2, 1–18, http://www. journalofvision.org/content/10/9/2, doi:10.1167/10.9.2. [PubMed] [Article]

Todd, J. T. (2004). The visual perception of 3D shape. *Trends in Cognitive Science, 8,* 116–121.

Todd, J. T., Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. L. (1996). Effect of changing viewing conditions on the perceived structure of smoothly curved surfaces. *Journal of Experimental Psychology: Human Perception and Performance, 22,* 695–706.

Tyler, C. W. (1998). Diffuse illumination as a default assumption for shape-from-shading in the absence of shadows. *Journal of Imaging Science and Technology, 42,* 319–325.

Waltz, D. (1975). Understanding line drawings of scenes with shadows. In P. H. Winston (Ed.), *The psychology of computer vision* (pp. 19–91). New York: McGraw-Hill.

Wright, M., & Ledgeway, T. (2004). Interaction between luminance gratings and disparity gratings. *Spatial Vision, 17,* 51–74.