The human visual system's assumption that light comes from above is weak

Yaniv Morgenstern, Richard F. Murray[1], and Laurence R. Harris

Department of Psychology and Centre for Vision Research, York University, Toronto, ON, Canada M3J 1P3

Edited by Dale Purves, Duke University Medical Center, Durham, NC, and approved June 14, 2011 (received for review January 19, 2011)

Every biological or artificial visual system faces the problem that images are highly ambiguous, in the sense that every image depicts an infinite number of possible 3D arrangements of shapes, surface colors, and light sources. When estimating 3D shape from shading, the human visual system partly resolves this ambiguity by relying on the light-from-above prior, an assumption that light comes from overhead. However, light comes from overhead only on average, and most images contain visual information that contradicts the light-from-above prior, such as shadows indicating oblique lighting. How does the human visual system perceive 3D shape when there are contradictions between what it assumes and what it sees? Here we show that the visual system combines the light-from-above prior with visual lighting cues using an efficient statistical strategy that assigns a weight to the prior and to the cues and finds a maximum-likelihood lighting direction estimate that is a compromise between the two. The prior receives surprisingly little weight and can be overridden by lighting cues that are barely perceptible. Thus, the light-from-above prior plays a much more limited role in shape perception than previously thought, and instead human vision relies heavily on lighting cues to recover 3D shape. These findings also support the notion that the visual system efficiently integrates priors with cues to solve the difficult problem of recovering 3D shape from 2D images.

3D shape perception | visual psychophysics | Bayesian modelling

Most people see Fig. 1 as a bas-relief footprint illuminated from the top of the page, even though it depicts a concave footprint illuminated from the bottom. This percept illustrates the light-from-above prior, the human visual system's implicit assumption that light shines from overhead (1–5). In most environments light originates above the horizon, so the light-from-above prior is a reasonable assumption that helps us choose the most probable interpretation of ambiguous images.

How does the human visual system resolve contradictions between the light-from-above prior and lighting cues in the many scenes where light does not shine from directly overhead? The evidence is inconclusive: some researchers have argued that visual lighting cues completely override the prior (6, 7), whereas others have maintained that lighting cues have either no effect at all on the perceived lighting direction that guides shape-from-shading (8–10), or less influence than nonvisual factors such as head orientation and the direction of gravity (11, 12). This question is pivotal, however, for understanding 3D shape perception. If the visual system relies heavily on the light-from-above prior instead of estimating lighting direction from visual cues, then human shape-from-shading mechanisms must not require accurate estimates of lighting direction (13) and hence differ profoundly from classic computer vision approaches to shape-from-shading (14).

To determine how the visual system resolves contradictions between the light-from-above prior and lighting cues, we probed shape-from-shading percepts under a range of lighting conditions. We showed observers ambiguously shaded disks embedded at random orientations in scenes where shading and shadow cues indicated the true direction of lighting (Fig. 2). The ambiguous disks could be interpreted as bumps illuminated from one direction or as dents illuminated from the opposite direction. Six observers judged whether a target disk in each scene looked like a bump or a dent. The lighting cues were sometimes strong (Fig. 2*A*) and sometimes weak (Fig. 2*B*). In separate blocks, light came from one of six evenly spaced directions (12 o'clock, 2 o'clock, 4 o'clock, and so on). In each condition (two lighting cue strengths × six lighting directions) we found the orientation at which disks looked most like bumps, and we took this to be the lighting direction that guided shape-from-shading processes in that condition (3, 15). We call this the "effective lighting direction." For example, if disks looked most bump-like when their brighter half was 30° clockwise of vertical, then we took the effective lighting direction to be 30° clockwise of vertical. To find the direction of each observer's light-from-above prior [which previous studies have found is not always exactly overhead (3)], we also measured the effective lighting direction in a block of trials in which the ambiguous disks appeared on flat circular surfaces that provided no lighting direction cues (Fig. 2*C*).

Results and Discussion

When the lighting cue direction θ_{cue} was the same as the light-from-above prior direction θ_{prior}, the effective lighting direction θ_{eff} was naturally the same as well, because the lighting cues simply reinforced the prior. However, we found that when the lighting cue direction shifted away from the prior direction, the effective lighting direction also shifted away. Fig. 3 plots the shift of the effective lighting direction away from the prior, $\theta_{eff} - \theta_{prior}$, as a function of the shift of the lighting cue direction away from the prior, $\theta_{cue} - \theta_{prior}$. Under strong lighting cues the effective lighting direction closely tracked the lighting cue direction (Fig. 3*A*), indicating that lighting cues almost completely overrode the prior. Intriguingly, under weak cues the effective lighting direction was neither the prior direction nor the cued direction, but instead was approximately halfway between the two (Fig. 3*B*). (Fig. S1 shows detailed data from a typical observer, and Fig. S2 shows the directions of individual observers' light-from-above priors.)

Evidently observers dealt with contradictions between the light-from-above prior and lighting cues by using an effective lighting direction that was a compromise between the two, and the compromise depended on how strong the lighting cues were. To examine this strategy more closely, we used a vector sum model of how observers combine information from two or more cues to estimate a direction (11, 16, 17). In this model the prior direction and the lighting cue direction are represented by unit vectors $\mathbf{v_{prior}}$ and $\mathbf{v_{cue}}$, respectively, and the effective lighting direction is a weighted sum of the two vectors, $\mathbf{v_{eff}} = w_{prior}\mathbf{v_{prior}} +$

Author contributions: Y.M., R.F.M., and L.R.H. designed research; Y.M. and R.F.M. performed research; Y.M. and R.F.M. analyzed data; and Y.M., R.F.M., and L.R.H. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

[1]To whom correspondence should be addressed. E-mail: rfm@yorku.ca.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1100794108/-/DCSupplemental.

Fig. 1. This photograph is usually seen as a raised footprint illuminated from above, even though it is actually an indented footprint illuminated from below. (Photograph courtesy of Manuel Cazzaniga.)

$w_{cue}\mathbf{v}_{cue}$. The weight ratio w_{prior}/w_{cue} determines whether the prior or the cues have a greater influence on the effective lighting direction. This model is well established in the literature on the subjective vertical (11, 17). We have recently shown that it is largely equivalent to a Bayesian cue combination model that assigns reliability weights to noisy directional cues [here the prior, which we treat as just another cue (18, 19), and the lighting cues] and combines the cues by making a maximum-likelihood direction estimate (16). Thus, observers who obey the vector sum model are following an efficient statistical cue combination strategy. In *SI Materials and Methods* we describe the vector sum model and our fitting methods in detail.

Fig. 3 shows the fit of the vector sum model. Under strong lighting cues the prior-to-cue weight ratio was $w_{prior}/w_{cue} = 0.13 \pm 0.04$ (maximum-likelihood fit and bootstrapped 95% confidence interval), confirming that strong cues had a much greater effect than the prior on the effective lighting direction ($w_{cue} > w_{prior}$), but also showing that the prior had a measurable residual influence ($w_{prior} > 0$). Under weak cues the weight ratio was $w_{prior}/w_{cue} = 1.10 \pm 0.14$, showing that even our weak cues had approximately as much influence as the prior ($w_{cue} \approx w_{prior}$). Because the prior and the weak cues had approximately equal influence, the weak-cue stimulus (Fig. 2*B*) can be seen as a visual representation of how much directional information observers received from the prior: clearly very little. The large spread in data points at lighting directions near ±180° in the weak condition is also accounted for by the vector sum model, which predicts highly variable direction estimates when the prior and the lighting cues have approximately equal weights but indicate opposite directions (16). In particular, although the fitted curve predicts that the effective lighting direction matches the prior direction when the prior and lighting cue directions are 180° apart, for some observers the effective lighting direction was actually much closer to the lighting cue direction. Fig. S3 and Table S1 report fits to individual observers' data, and we discuss individual differences further in *SI Discussion*.

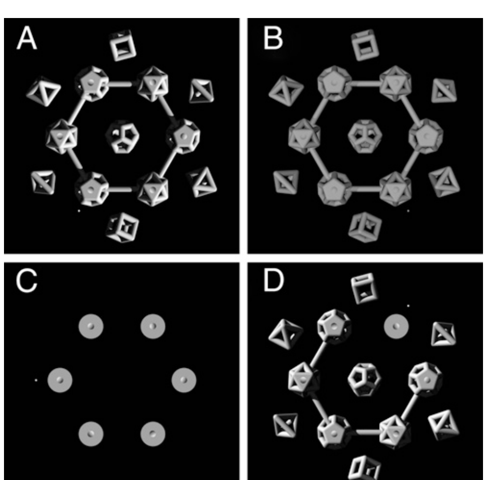

Fig. 2. Typical stimuli. Stimuli in (*A*) strong cue, (*B*) weak cue, (*C*) no cue, and (*D*) no local cue conditions. Observers judged whether the shaded disk next to the small white dot looked like a bump or a dent. Here the lighting directions are (*A*) 4 o'clock, (*B*) 8 o'clock, and (*D*) 10 o'clock.

As an independent measure of lighting cue strength, we measured how precisely observers could rotate an on-screen arrow to indicate the lighting direction in the strong and weak cue conditions. We found that observers' angular errors had circular SD 33° under strong lighting cues and 66° under weak cues (20). For comparison, the circular SD of responses randomly distributed within ±90° of the correct direction is 54°. The high error for judging lighting direction using the weak cues shows that these cues were almost unusable, making it all the more remarkable that they were even partially able to override the light-from-above prior. [Explicit estimates of lighting direction depend on the prior and lighting cues in a manner consistent with our model of the effective lighting direction that guides shape-from-shading (21), but it is nevertheless possible that our observers' explicit estimates of lighting direction differed from the implicit estimates that guided their responses in the main experiment (22).]

We found that observers assigned approximately as much weight to the weak cue stimulus as to the light-from-above prior, and we also found that observers' explicit lighting direction estimates based on the weak cue stimulus had a circular SD of 66°. This suggests that 66° can be taken as a rough estimate of the width of the light-from-above prior. This value is similar to the spread of lighting directions that an observer may encounter over the course of the day: the circular SD of directions uniformly distributed over the upper semicircle is 54°, as is the circular SD of lighting directions uniformly distributed over the upper hemisphere and projected into the frontoparallel plane.

Do these findings mean that the light-from-above prior plays little role in shape-from-shading in real-world scenes that are rich with lighting cues? One caveat is that the lighting cues in our experiment were directly adjacent to the ambiguous disks. In real scenes lighting can vary from place to place, so perhaps the visual system only allows lighting cues to affect the perceived shape of immediately adjacent objects (23) and relies on the light-from-above prior for interpreting large regions that have no local lighting cues. To test this possibility, we eliminated local lighting cues in the strong cue stimulus by showing the target disk on a flat surface and removing the two rods adjacent to the target disk (Fig. 2*D*). Eliminating local lighting direction cues had little effect: the effective lighting direction at the target disk still closely tracked the direction of the lighting cues, even though the nearest cues were on separate objects at least 3.3° away ($w_{prior}/w_{cue} = 0.00 \pm 0.03$; Fig. 3*C*).

Conclusion

The finding that the role of lighting direction cues in shape-from-shading depends heavily on lighting cue strength may explain the inconsistent conclusions of previous studies. Studies that used complex illuminated objects with strong lighting direction cues concluded that lighting cues guide shape-from-shading (6, 7), whereas studies that used weaker lighting cues concluded that they have little effect (8–12). We discuss the lighting cues in previous studies in further detail in *SI Discussion*.

Given the attention paid to the light-from-above prior in previous literature (1–12, 15, 24, 25), one would think that it played a crucial role in shape perception. In fact, the light-from-above prior has a surprisingly weak influence and is easily overridden. Using a weak prior is a rational strategy for the visual system to follow in a world where knowing the current lighting direction is important and where, on average, light comes from overhead, but

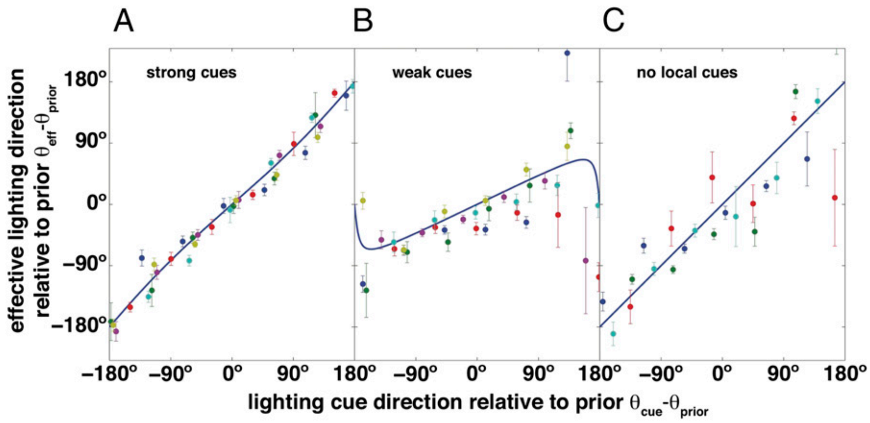

Fig. 3. Effective lighting direction as a function of lighting cue direction. Panels correspond to (*A*) strong cue (six observers), (*B*) weak cue (six observers), and (*C*) no local cue conditions (four observers). Angles are measured relative to the direction of each observer's individually determined light-from-above prior. Different colors correspond to different observers. Error bars indicate SE. Solid lines are maximum-likelihood fits of the vector sum model.

where there are large variations in lighting direction that are reliably cued by shading and shadows.

Materials and Methods

Stimuli and Apparatus. The stimuli were computer-generated images of matte objects rendered in RADIANCE (26). The simulated lighting consisted of two distant point sources. One source was in one of six evenly spaced directions at 12 o'clock, 2 o'clock, 4 o'clock, and so on, 30° toward the viewer from the frontoparallel plane (like clock hands bent 30° forward from the clock face). The second source was in the direction of the viewer. In the strong cue condition the first source was brighter, and in the weak cue condition the second was brighter. In both cases the brighter source contributed 85% of the illuminance of the frontoparallel planes that the ambiguous disks appeared on, which was always 120 cd/m^2. The no cue stimulus (Fig. 2*C*) showed ambiguous disks on six circles of diameter 2.3° and luminance 120 cd/m^2. The ambiguous disks had diameter 0.63°, and the scene subtended 17.4° horizontally.

Procedure. In each condition (strong cue, weak cue, no local cue), observers participated in seven blocks. Six blocks showed scenes with six lighting directions, and the seventh showed scenes without lighting cues (Fig. 2*C*). On each trial, six ambiguous disks appeared at various orientations, with a white dot next to the target disk (e.g., lower left in Fig. 2*A*). The observer judged whether the target disk looked like a bump or a dent. We showed six disks because observers found the task easier if they could compare the target disk with other disks.

Analysis. We made a maximum-likelihood fit of a periodic function to the probability of a "bump" response as a function of the orientation of the target disk, for each observer in each condition. The effective lighting direction was the orientation where the fitted curve peaked. Fig. 3 shows fits of the following equation, which follows from the vector sum model:

$$\theta_{eff}-\theta_{prior} = \mathrm{arctan2}\left(\sin(\theta_{cue}-\theta_{prior}),\ (w_{prior}/w_{cue})+\cos(\theta_{cue}-\theta_{prior})\right). \quad [1]$$

Here *arctan2* is the four-quadrant inverse tangent. See *SI Materials and Methods* for further details.

ACKNOWLEDGMENTS. We thank Minjung Kim and two anonymous reviewers for comments on the manuscript. This work was funded by grants from the Natural Sciences and Engineering Research Council and the Canada Foundation for Innovation.

1. Metzger W (2006) *Laws of Seeing* (MIT Press, Cambridge, MA).
2. Ramachandran VS (1988) Perception of shape from shading. *Nature* 331:163–166.
3. Sun J, Perona P (1998) Where is the sun? *Nat Neurosci* 1:183–184.
4. Mamassian P, Goutcher R (2001) Prior knowledge on the illumination position. *Cognition* 81:B1–B9.
5. Adams WJ (2007) A common light-prior for visual search, shape, and reflectance judgments. *J Vis* 7(11):11, 1–7.
6. Berbaum K, Bever T, Chung CS (1984) Extending the perception of shape from known to unknown shading. *Perception* 13:479–488.
7. Ramachandran VS (1988) Perceiving shape from shading. *Sci Am* 259:76–83.
8. Erens RGF, Kappers AML, Koenderink JJ (1993) Estimating local shape from shading in the presence of global shading. *Percept Psychophys* 54:334–342.
9. Mingolla E, Todd JT (1986) Perception of solid shape from shading. *Biol Cybern* 53: 137–151.
10. Wagemans J, van Doorn AJ, Koenderink JJ (2010) The shading cue in context. *i-Perception* 1:159–178.
11. Jenkin HL, Jenkin MR, Dyde RT, Harris LR (2004) Shape-from-shading depends on visual, gravitational, and body-orientation cues. *Perception* 33:1453–1461.
12. Yonas A, Kuskowski M, Sternfels S (1979) The role of frames of reference in the development of responsiveness to shading information. *Child Dev* 50:495–500.
13. Koenderink JJ, van Doorn AJ (1980) Photometric invariants related to solid shape. *Opt Acta (Lond)* 27:981–996.
14. Zhang R, Tsai PS, Cryer JE, Shah M (1999) Shape from shading: A survey. *IEEE T Pattern Anal* 21:690–706.
15. Hess EH (1961) Shadows and depth perception. *Sci Am* 204:138–148.
16. Murray RF, Morgenstern Y (2010) Cue combination on the circle and the sphere. *J Vis* 10(11):15, 1–11.
17. Mittelstaedt H (1983) A new solution to the problem of the subjective vertical. *Naturwissenschaften* 70:272–281.
18. Mamassian P, Landy MS (2001) Interaction of visual prior constraints. *Vision Res* 41: 2653–2668.
19. Backus BT (2009) The mixture of Bernoulli experts: A theory to quantify reliance on cues in dichotomous perceptual decisions. *J Vis* 9(1):6, 1–19.
20. Fisher NI (1993) *Statistical Analysis of Circular Data* (Cambridge Univ Press, Cambridge, UK).
21. O'Shea JP, Agrawala M, Banks MS (2010) The influence of shape cues on the perception of lighting direction. *J Vis* 10(12):21, 1–21.
22. Hillis JM, Brainard DH (2007) Distinct mechanisms mediate visual detection and identification. *Curr Biol* 17:1714–1719.
23. Ostrovsky Y, Cavanagh P, Sinha P (2005) Perceiving illumination inconsistencies in scenes. *Perception* 34:1301–1314.
24. Rittenhouse D (1786) Explanation of an optical deception. *T Am Philos Soc* 2:37–42.
25. Brewster D (1826) On the optical illusion of the conversion of cameos into intaglios, and of intaglios into cameos, with an account of other analogous phenomena. *Edinb J Sci* 4:99–108.
26. Ward LG (1994) The RADIANCE lighting simulation and rendering system. *Proc SIGGRAPH '94*, ed, Glassner A (Association for Computing Machinery, New York), pp 459–472.