Max–Planck–Institut
für biologische Kybernetik
Spemannstraße 38 • 72076 Tübingen • Germany

Technical Report No. 73

Perception of shape from shading on a cloudy day

Michael S. Langer and Heinrich H. Bülthoff

October 1999

[1] AG Bülthoff, E–mail: michael.langer@tuebingen.mpg.de

This document is available as /pub/mpi-memos/TR-73.ps via anonymous ftp from ftp.kyb.tuebingen.mpg.de or from the World Wide Web, http://www.kyb.tuebingen.mpg.de/bu.html.

Perception of shape from shading on a cloudy day

Michael S. Langer and Heinrich H. Bülthoff

Abstract. The human visual system has a remarkable ability to interpret smooth patterns of light and shade on a surface in terms of 3-D surface geometry. Classical studies of shape–from–shading have assumed that surface luminance depends on the local surface orientation. This classical shading model holds, for example, on a sunny day. A common situation in which the classical model fails to hold, however, is a diffuse lighting condition such as on a cloudy day. Here we report on the first set of psychophysical experiments that explicitly address perception of shape–from–shading under diffuse lighting. Our main findings are that depth discrimination under diffuse lighting is superior to that predicted by a classical sunny day model, and by a model in which depth varies with perceived luminance *i.e.* dark means deep . We find that performance is correlated with the latter model, however, under both diffuse source and point source lighting. The results imply that the visual system uses multiple models when perceiving shape–from–shading.

1 Introduction

For centuries, artists have studied how patterns of light and shade on a surface can provide visual cues to surface shape. Shading is now considered to be a fundamental visual cue, along with binocular disparity, texture, and contour. While it is widely accepted that the visual system uses shading information to perceive surface shape, it is not yet known about how the visual system does so and how well. Previous studies of shape-from-shading perception have concentrated on one shading model only, in which image intensity varies with the local surface orientation. This model is valid, for example, when there is a well–defined point light source direction such as on a sunny day (Fig. 1). In this paper, we address shape–from–shading perception under an alternative lighting condition that is very common in nature, namely diffuse lighting such as on a cloudy day (Fig. 2) and compare performance under diffuse lighting to that under standard sunny day conditions. We begin the paper by discussing how the appearance of a surface differs under a sunny day lighting condition versus a diffuse lighting condition.

1.1 Shading on a sunny day

The classical model of shading is that the luminance of a surface element depends on the orientation of that element with respect to the light source direction. Surface elements that face the source directly are brighter than those that have an oblique orientation with respect to the source because the latter elements are foreshortened with respect to the source, and hence do not receive as much light per unit area.

Formally, this classical model of shading may be stated as follows. Let the image coordinates of a point on the surface be (x, y) and assume the surface is viewed under orthographic projection. Let the height function of the surface with respect to the optical axis be $z(x, y)$. Let $\mathbf{N}(x, y)$ be the unit surface normal. Let the surface have uniform Lambertian reflectance. Let the light source direction be $\mathbf{L}$. Then, assuming the point light source is visible from the surface at (x, y), the image intensity at (x, y) may be modeled as

$$\mathrm{I}(x, y) = \mathbf{N}(x, y) \cdot \mathbf{L}\ . \qquad (1)$$

Several qualifiers to this classical model should be stated. First, no real surface has exactly Lambertian reflectance, as the luminance of a point typically does depend on the angle from which the point is viewed. Second, even if a scene does have a well–defined light source direction such as the sun, there is still a diffuse component to the illumination that must be considered. For example, on a sunny day, one cannot ignore the illumination from the blue sky. Third, surface elements are illuminated not only by light sources but also by each other via interreflections. Given these qualifiers, one must regard the above model merely as a "rough and ready" approximation to surface appearance on a sunny day.

1.2 Shading on a cloudy day

The appearance of a surface under diffuse lighting such as on a cloudy day is very different from its appearance under point source lighting. Under diffuse lighting, surface luminance depends primarily on the amount of the source that is visible from the surface. Intuitively, hills of a surface tend to be brighter than valleys since more of the diffuse source is visible from the hills. This effect is apparent, for example, in human faces. Wrinkles in the skin tend to be yields dark lines, since the less of the diffuse source is visible from within the concave part of the fold. Similarly, nostrils tend to also be dark. Note that this is really a shadowing effect since the walls of a concavity limit the amount of the diffuse source that is visible from within the concavity.

Formally, we can model shading under diffuse lighting by adapting the model of Eq. (1) to the diffuse condition. We assume Lambertian reflectance and ignore interreflections, and we let the diffuse source be uniform over all directions and have unit radiance. Formally, let $\theta(x, y)$ be the solid angle of the diffuse source that is visible from surface point (x, y). Let $d\Omega$ denote an infinitesimal angle centered at a direction $\mathbf{L} = (-l_x, -l_y, l_z)$. Image intensity at (x, y) may be modeled by integrating over $\theta(x, y)$,

$$\mathrm{I}(x, y) \;=\; \frac{1}{\pi} \int_{\theta(x,y)} \mathbf{N}(x, y) \cdot \mathbf{L}\ d\Omega\ . \qquad (2)$$

The key difference between Eqs. (1) and (2) lies in the effects of shadowing. In Eq. (1), shadowing is a binary phenomenon: a point either lies in shadow or it does not[1]. In Eq. (2), shadowing is continuous. Except for the highest hilltops, all points lie partly in shadow and the amount of shadowing varies continuously from one point to another.

1.3 Shape–from–shading psychophysics

How can we model human perception of shape from shading? Most studies in the past have centered around two themes. The first concerns ambiguities in the relationship between shading and shape, and how the visual system resolves these ambiguities. The second theme concerns the accuracy and consistency with which observers judge surface shape. Let us briefly the discuss these two themes in turn.

The main ambiguity between shape and shading that has been addressed in the past is that, under point source lighting, shading and shape have a depth–reversal symmetry: a surface illuminated from one direction produces the same retinal image as a depth–reversed surface illuminated from a mirror symmetric direction. This depth–reversal symmetry arises directly from Eq. (1), and is illustrated qualitatively in Figure 1a. Formally, the ambiguity arises because for a given surface $z(x, y)$ and a given lighting direction $\mathbf{L}$, the same image $\mathrm{I}(x, y)$ is produced when a depth–reversed surface, $-z(x, y)$ is illuminated from a mirror symmetric direction, $\bar{\mathbf{L}} = (-l_x, -l_y, l_z)$. This

[1] For a small compact source such as the sun, a penumbra is produced where which the source is partially visible.

ambiguity has been discussed in the literature for over two hundred years (Rittenhouse, 1786; Brewster, 1826).

The visual system can resolve this ambiguity using several strategies. For example, it can uses cues other than shading. In the rendered image of Fig. 1a, there are global shape cues that imply that the surface has an overall convexity, rather than an overall concavity. These include perspective cues, as well occluding contours cues on the flanking regions of the surface. Another strategy to resolve the depth–reversal ambiguity is to invoke prior knowledge about typical environments. The visual system tends to assume a light source from above, rather than a light source from below. That is, bright points are seen as having a surface normal that points upwards (above the line of sight), and dark points are seen as having a surface normal that points downwards (below the line of sight). This preference is ecologically valid since the sun is typically above the line of sight. (See (Berbaum, Bever, & Chung, 1983, 1984; Ramachandran, 1988; Howard, Bergstrom, & Ohmi, 1990) for papers related to this idea). There is also evidence that a source from the left is preferred over one from the right (Sun & Perona, 1998), although this effect is smaller than the above vs. below effect. Preferences on shape have also been found. The visual system prefers surfaces that tilt upwards as a floor rather than downwards as a ceiling (Reichel & Todd, 1990), and it prefers familiar shapes such as faces over unfamiliar ones such as hollow masks (Gregory, 1970).

The second important theme in shape–from–shading perception has concerned the accuracy and consistency to which a shape is perceived from shading. Many shape properties have been psychophysically measured including the relative depth of point pairs, the surface orientation (slant and tilt), and the surface curvature. Examples of such studies include (Bülthoff & Mallot, 1988; Johnston & Passmore, 1994a, 1994b; Todd & Mingolla, 1983; Mingolla & Todd, 1986; Mamassian, Kersten, & Knill, 1996; Mamassian & Kersten, 1996; Norman & Todd, 1996; Koenderink, van Doorn, & Kappers, 1992; Reichel, Todd, & Yilmaz, 1995; Koenderink, van Doorn, & Kappers, 1996).

1.4 Shape-from-shading in computer vision

The computational approach to shape–from–shading has a long history. Ernst Mach was the first to formulate shape–from–shading as an abstract computational problem over a century ago (Mach, 1866). His formulation was based on the point source model of Eq. (1). The first computer vision algorithm for solving this problem appeared in the late 1960's and is due to Horn (Horn, 1975). Many algorithms have been presented since. For example, see the collection of papers in (Horn & Brooks, 1989) as well as (Pentland, 1990; Dupuis & Oliensis, 1994).

The diffuse lighting version of shape–from–shading has received less attention historically. In computer vision, it was first solved by Langer and Zucker in (Langer & Zucker, 1994) who presented an algorithm based on an approximation of Eq. (2) in which image intensity depends only on angle $\theta(x,y)$ of the visible source. A more accurate algorithm was presented by Stewart and Langer (Stewart & Langer, 1997) that accounted for surface orientation effects and interreflections as well.

1.5 The gap between human perception and computer vision

Although the computational problems of shape–from–shading as defined by Eqs. (1) and (2) are now well understood, and although much has been learned from psychophysics about how shape is perceived from shading, the link between computer vision and human vision has been difficult to make. One reason is that, in computer vision, the models of Eq. (1) and (2) are typically treated as if they are correct for a given situation, and researchers have concentrated on the numerical issues of computing a precise solution, given the model. For example, one typically assumes that the lighting conditions are known or can be accurately estimated (Pentland, 1982). This assumption is difficult to justify for human vision, especially in a realistic lighting conditions which typically fall between the two ideals of Eq. (1) and (2).

A second difficulty in relating computer vision to human vision is that, in human vision, it is non-trivial to measure a perceived surface. Indeed, when one considers issues of non–uniformity of spatial vision, attention, eye movements, and memory, the term "the perceived surface" becomes somewhat nebulous. Even if one psychophysically measures local surface normals and fits these normals to a surface, one cannot conclude that the computed surface is the surface perceived by the observer (see (Koenderink et al., 1992) for a nice discussion of this point). Contrast this difficulty with the situation in computer vision where an algorithm computes a precise representation of an entire surface height function and one can evaluate the algorithm immediately by comparing the computed height to ground truth.

On a more positive note, we argue that computer vision models can still offer much insight into perception, despite the above difficulties. Computer vision models provide a theoretical background from which hypothesis about human vision can be generated. In the present paper, we use insights obtained recently in computer vision from studies of shape–from–shading under diffuse lighting to test how well humans perceive shape–from–shading under this condition.

1.6 Overview

Our main goal in the present paper is to measure perceived shape from shading under diffuse lighting and point source lighting, and to understand how differences in performance in these lighting conditions are related to a particular task, *i.e.* computational problem, that an observer is asked to solve. There are three experiments. The first (Sec. 2) establishes that a set of shading stimuli give rise to shape perception, and that observers show differences between lighting conditions even for a basic task of judging whether a single point is convex or concave. The second (Sec. 3) tests the hypothesis that observers perceive shape–from–shading by relating surface height directly to luminance. Such a strategy is plausible since height is statistically correlated with luminance under diffuse lighting. The results show that observers perform better than if they were using this strategy. The third experiment (Sec. 4) is a control on the second, and attempts to tease apart the different strategies that the visual system might be using. Finally, in Sec. 5, we compare observer performance to that of various computer vision models that have been proposed for the diffuse lighting condition.

2 Experiment 1: qualitative shape

The first experiment compares qualitative shape perception under point source and diffuse source conditions. Observers were asked to judge whether isolated marked points on a surface are "on a hill" or "in a valley." The main purpose of the experiment was to establish any basic differences between point source and diffuse conditions.

In pilot studies, we used shaded images of smooth terrain surfaces that were fronto–parallel with respect to the line of sight. We found that naive observers typically failed to perceive shape–from–shading for such surfaces, either under under point source or diffuse lighting. For this reason, we decided to use rendered surfaces that have a more complex shape. Two examples of the surfaces we used are shown in Figs. 1b and 2b. The reader may verify that these images yield a vivid impression of shape.

We believe that the vivid impression of shape for these surfaces relative to the terrain surfaces is that the former contain several other cues besides shading, and these cues provide a scene context in which the shading information can be interpreted.

These cues included two disk frames seen in perspective, as well as occluding contours on the surface near the flanking region of stimulus. Occluding contours in particular are believed to provide a strong cue to shape (Howard, 1983; Todd & Reichel, 1989; Ikeuchi & Horn, 1981; Koenderink, 1984; Todd & Mingolla, 1983; Mamassian & Kersten, 1996). To limit the use of the occluding contours information, we tested perceived shape only near the central region of each stimulus. In this region, the surface was essentially a fronto–parallel terrain and no occluding contours were present. We also used relatively short presentation times. A more detailed discussion of the method now follows.

2.1 Method

2.1.1 Stimuli

Surface shapes were defined by modulating the radius of a cylinder with low pass filtered white noise. The cylinder was defined by a 1024×1024 polygonal mesh and the cutoff frequency was 60 cycles. Each cylinder was then sliced into eight disks of size 1024×128. Each of these disks was rendered from four different viewing directions, separated by 90 degree increments.

The disk surfaces were rendered using the RADIANCE software package (Larson & Shakespeare, 1998). Surface material was Lambertian with reflectance of 30 per cent and interreflections were computed to two bounces. Five lighting conditions were used: a uniform diffuse condition, and four different point source conditions (above–left, above–right, below–left, below–right). In the diffuse condition, the source was a uniform sphere surrounding the object. In each point source conditions, the source was 15 degrees from the line of sight. This angle was small enough that no cast shadows appeared in the central test region. A weak diffuse source component was added in each point source condition to simulate the ambient illumination.

For the diffuse lighting images, the renderings contained a small amount of pixel noise which was due to the stochastic sampling of the source. For the rendering parameters we used, the noise was rather small. It was roughly 1 per cent (root mean square error) of the maximum intensity. This error was independent of the intensity of a pixel. Hence for darker pixels the relative error was larger. [2] Noise of roughly the same order of magnitude was introduced by the interreflection term.

Images were presented on a CRT monitor which was calibrated so that screen luminance was linearly related to rendered surface luminance. Intensities in each image were normalized so that all images had the same maximum intensity. Observers wore an eye patch over the non-dominant eye and viewed the stimuli in a dimmed room at a distance of 2 m. Each surface subtended a viewing angle of 10×4 degrees.

Probe points were chosen from the central 2×2 degree square. Probe points were chosen only if they had positive Gaussian curvature, that is, saddle points were not used.

2.1.2 Observers

Twenty observers participated (age 18–30). All had normal or corrected–to–normal vision.

2.1.3 Procedure

Each trial consisted of a grey silhouette that was presented for 0.2 s, followed by a probe for 0.8 s during which time the observer made an eye movement to the probe. A stimulus image was then presented with the probe superimposed on the image. The probe was 6 pixels square prior to the eye movement, and 4 pixels square when it was superimposed on the image (as in Fig. 1b).

The observers' task was to judge whether the probe was "on a hill" or "in a valley" (see Fig. 1b). The observer had a maximum of 1.2 s to respond by pressing on a left or

[2] We estimated the noise in the rendering by using a set of spherical concavities of different shapes and by comparing the rendered images to an analytic expression of the ground truth (Stewart & Langer, 1997).

right response key. No feedback was given. Responses longer than 1.2 s were discarded and the same trial condition was repeated at the end of the same block.

Lighting conditions were balanced over all trials. Half of the observers ran a mixed condition in which the lighting varied randomly from trial to trial, and half ran a blocked condition in which the lighting condition was constant within each block and varied between blocks. (No effect was found for mixed vs. blocked conditions, and the data were subsequently pooled. We present only the pooled data.)

Each observer ran 320 trials. These were divided into 10 blocks of 32 trials each. Prior to the experiment, each observer ran a practice session of a single block. No feedback was given.

2.2 Results

Percent correct scores are shown in Fig. 3a. In the point source conditions, percent correct was higher when the source was above the line of sight than when it was below the line of sight ($F(1,18) = 86, p < .0001$). This replicates the classical finding that the visual system prefers light from above *i.e.* bright points are perceived as having an upward orientation and darker points are perceived as having a downward orientation. Percent correct was higher when the source was from the left than from the right (Sun & Perona, 1998), but this effect was not significant ($F(1,18) = 2.2, p = .15$).

In the diffuse condition, percent correct was well above chance ($t(19) = 11.6, p < .0001$) and was as high as under the best point source tested.

2.3 Discussion

If the visual system used a point-light-source-from-above model to perceive shape–from–shading, then performance in the diffuse lighting condition would have been at chance. Under diffuse condition, an equal amount of light arrives at a given surface from above the line of sight as from below, and so a point–source–from–above model must yield the same level of performance as a point–source–from–below model. But these two yield opposite responses (hill vs. valley) on each trial because of a depth reversal symmetry under point source lighting. It follows that each model would yield chance performance under a diffuse lighting condition. Since observers were well above chance under the diffuse condition, an alternative model must be considered.

One alternative model that could account for the performance in the diffuse condition is suggested by a heuristic used in artistic rendering, in which depth is conveyed by darkness (Nicolaides, 1941). This phenomenon that dark means deep has been known for centuries. For example, Leonardo da Vinci wrote that "among bodies equal in size and distance, that which shines the more brightly seems to the eye nearer" (E. MacCurdy, 1938). Such a dark–means–deep model has been formally shown to bias shape perception in studies of line drawings (Dosher, Sperling, & Wurst, 1986) and in point source shading (Christou & Koenderink, 1997; Bülthoff & Mallot, 1988). It has been hypothesized that this model is used as a default assumption for shape–from–shading under diffuse lighting as well (Langer & Zucker, 1994; Tyler, 1998), although this hypothesis has not been formally tested. Our second experiment was therefore intended to hypothesis that the visual system uses a dark–means–deep model to infer shape–from–shading under diffuse lighting.

3 Experiment 2: depth discrimination

One reason why we might not expect a dark–means–deep model to be used under diffuse lighting is that, strictly speaking, this model is incorrect. Although surface height and image intensity are correlated under diffuse lighting, the correlation is not perfect. Under diffuse lighting, a valley may contain a local intensity maximum at its deepest point (Fig. 2a). This local maximum may have several causes. The surface normal may turn toward the vertical, and hence towards the visible part of the source yielding a classical $\mathbf{N} \cdot \mathbf{L}$ shading effect. Alternatively, there may be a local maximum in the visible angle of the source. (A third possibility, which can occur for non–Lambertian surfaces hence not for the surfaces we use, is that there may be a specular reflection of the diffuse source from the valley.) If the visual system were to use a dark–means–deep model to perceive shape–from–shading under diffuse lighting, then such a local intensity maximum in a valley should yield an illusory local hill. Our second experiment tested for this illusion.

The second experiment used the same stimuli as the first. In each trial of the second experiment, a pair of points on the surface was marked, and the task now was to judge which of the two points was higher, that is, closer to the observer along the line of sight (Koenderink et al., 1996). The idea of the experiment is to create a conflict situation in which the two points have different intensities and different heights. This yields the two conditions shown in Fig. 3b : either the intensity and height differences are of the same sign (**correlated**) or they are of opposite sign (**anti–correlated**). If the visual system were to rely on dark–means–deep as a cue to shape, then there should be systematic differences in performances under these two conditions. Details of the method are as follows.

3.1 Method

3.1.1 Stimuli

For each surface, pairs of points were chosen such that depth varied monotonically between the two points (Todd & Reichel, 1989). Moreover, both the image intensities and the heights of the two points differed by a given amount. Let I_1, I_2 and z_1, z_2 be the intensities and heights of two points, respectively. Each probe pair was required to have an image contrast $|\frac{I_1 - I_2}{I_1 + I_2}|$ that was 0.6 to 0.8 times the standard deviation of image contrast over the central test square of the image, and similarly, each pair was required to have a height difference $|z_1 - z_2|$ that was 0.6 to 0.8 the standard deviation of height differences of the central test square.

Correlated and anti–correlated conditions were balanced within each lighting condition. In this way, a dark–means–deep model predicted chance performance within a given lighting condition.

3.1.2 Procedure

Apart from the difference in the task, the procedure was the same as the first experiment. The only other difference is that there were 384 trials partitioned into 12 blocks of 32 trials each, and only two point source conditions were used, namely above–left and below–right.

3.1.3 Observers

Seventeen observers participated. Each was naive to the purpose of the experiment and none participated in any other experiment.

3.2 Results

The results are shown in Fig. 3b. In the point source conditions, percent correct was higher when the source was to the upper–left than to the lower–right ($F(1, 15) = 88.7$, $p < .0001$), replicating the classical finding that the visual system prefers light from above rather than below. In particular, when the source was below the line of sight, observers were below chance, implying that they were systematically fooled by the depth–reversal ambiguity.

In the diffuse condition, percent correct was higher in the correlated condition than in the anti–correlated condition. Moreover, overall percent correct in the diffuse condition was 63 percent which is significantly above chance (paired t-test, $t(16) = 8.1, p < .0001$).

3.3 Discussion

How can we account for the performance in the diffuse condition? One possibility is that observers were not able to resolve the shading as well in the anti–correlated condition as in the correlated condition. This may have led to a greater amount of guessing in the anti–correlated condition, and hence to performance that more near to chance. The shading in the anti–correlated condition may indeed have been more difficult to resolve, for the following reason. The anti–correlated trials tended to occur within valleys, while the correlated trials tended to occur on hills (see Fig. 2). Hence points in the anti–correlated trials were darker on average than points in the correlated trials. We tried to control for this systematic luminance difference between conditions, by using contrast rather than luminance when choosing probe pairs (see Sec. 3.1.1). However, for complex shaded images, the Michelson contrast measure that we we used may not have exactly captured perceived contrast (Peli, 1990). Hence the effect we found in the diffuse condition could, at is stands, be explained just in terms of contrast detect ability. A further experiment will test this issue.

An alternative account of the above–chance performance in the diffuse condition is that observers were indeed able to resolve the contrast difference in the anti–correlated condition, but that they nonetheless correctly perceived the brighter point to be deeper. If this is the case, it would be remarkable, and it would raise the question of what perceptual model could best describe this behavior.

A third possibility is that observers may have used a dark–means–deep model on some trials, but may have guessed on other trials. Such a strategy would explain why performance was better in the correlated condition than in the anti–correlated condition. However, it would not explain why performance was above–chance overall, since both guessing and dark–means–deep strategy would each yield chance performance overall. Thus, a dark–means–deep model may play a role, but it is not sufficient to account for the above chance performance.

4 Experiment 3: brightness discrimination

To try to disentangle the three factors just discussed, we carried out a third experiment. This experiment was identical to the second experiment, except that now the task was to discriminate the brightness of pairs of points, rather than the depth. That is, the task was to judge which of the two surface points was brighter, rather than which was higher. The main idea is that if observers had used a dark–means–deep model in the depth discrimination task, then the behavior in that depth task should be the same as in a brightness discrimination task.

4.1 Method

Eleven new observers participated. Each ran the mixed condition in which the lighting condition varied randomly from trial to trial. The experiment was otherwise identical to Experiment 2, except that the task was to judge which point was brighter rather than which was higher.

4.2 Results

The results are shown Figure 3b. We first consider the diffuse lighting condition. If observers had used a dark–means–deep strategy in the depth task (Experiment 2), then the behavior in the depth task should have been the same as that in the brightness task (Experiment 3). This was not the case, however. Responses in the depth task were significantly different than responses in the brightness task ($F(1,26) = 16.8, p < .001$). In particular, in the anti–correlated condition, observers were 77 percent correct in the brightness task but only 53 percent “correct” in depth task. (The quotation marks on “correct” emphasize the null hypothesis that observers were using a dark–means–deep strategy.) This difference is a significant ($F(1,26) = 27, p < .0001$) and implies that observers did not rely entirely on a dark–means–deep strategy in the depth task.

A dark–means–deep strategy did appear to play a role, however. A *post hoc* trial–

by–trial comparison of the responses in the anti–correlated condition of the depth and brightness tasks revealed a significant correlation ($r = .544, p < 0.0001$). That is, although observers were roughly at chance in the anti–correlated condition of the depth task, they were not merely guessing. Rather, certain trials tended to yield either a correct response or an incorrect response, and this tendency was consistent with a dark–means–deep strategy. The trial–to–trial correlation was slightly larger when the correlation was computed over the correlated and anti–correlated trials.

We also found a significant trial-to-trial correlation for the sunny day conditions, both for the point source above–left ($r = .21, p = .017$) and for the point source below–right ($r = .28, p = .011$). Again, this correlation does not fully account for the observers' behavior under the sunny day conditions, as they were well above chance when the light source was from above and well below chance when the light source was from below. The result does suggest, however, that a dark–means–deep model does play a role (albeit a secondary one) under the point source conditions (Christou & Koenderink, 1997), just as it did under the diffuse condition above.

Finally, we note that in the brightness task and under diffuse lighting, performance was higher in the correlated condition than in the anti–correlated condition ($F(1,10) = 19, p < .001$). Whether this difference was due to the complexity of brightness perception in complex images (Peli, 1990) or whether it was due to differences in the rendering noise between conditions, we cannot say. Either way, the effect is small in comparison to the difference in performance between the brightness and depth tasks. Hence, the effect does not account for the main finding, which is that observers performed better under the diffuse condition than if they had used a dark–means–deep strategy.

5 General Discussion

We have shown that a dark–means–deep model is not sufficient to explain the performance under diffuse lighting in the depth discrimination task. We have also shown that observers had greater difficulty in the anti–correlated condition than in the correlated condition in that task. Was this difficulty merely due to a bias to perceive dark points as deep? Or was the anti–correlated condition inherently more difficult than the correlated condition ? To try to address these questions, we compared performance of the observers in the diffuse condition to that of various computational models.

5.1 Computational Models

We first considered the performance of two computer vision algorithms (Langer & Zucker, 1994; Stewart & Langer, 1997) at the depth discrimination task. In each trial, we ran each algorithm on a 2 degree square neighborhood in the image. The probe pair was in the center of this neighborhood. The depth discrimination judgment was based directly on the shape computed by each algorithm.

The percent correct scores are shown in Table 1 for the correlated and anti–correlated conditions, along with trial-by-trial correlations. Three points are worth noting. First, the SL algorithm (Stewart & Langer, 1997) has a less than perfect score in the anti–correlated condition, even though this algorithm is given the correct shading model used to render the images. This suggests that human performance may be partly limited by the inherent difficulty of the task, especially in the anti–correlated condition. Second, the LZ algorithm (Langer & Zucker, 1994) was below chance in the anti–correlated condition. This implies that the surface normal effects are important for accurately judging shape under diffuse lighting, especially for judging the shape of valleys which contain local intensity maxima. Third, although the trend in the performance is similar between the computational

	corr	anti	r
LZ	100	27	.13
SL	100	77	$-.03$
blur	80	47	.07

Table 1: Percent correct scores and trial–to–trial correlation.

models and the human observers (namely better performance in the correlated condition than in the anti–correlated condition and above–chance performance overall), there was no trial–to–trial correlation between the human observers and the algorithms (third column in Table 1). Thus, we cannot explain observers' performance in this task in terms of the SL and LZ algorithms.

A third computational model to consider is that observers somehow "ignored" the local intensity maxima in valleys. As we mentioned earlier, probe pairs in the anti–correlated trials tended to be darker than those in the correlated trials, and observers may have opted for a dark–means–deep strategy in the former case. One way to implement such a strategy would be to blur the image slightly, prior to applying a dark–means–deep model. Such blurring would remove the small local intensity maxima within valleys. We found that by using a Gaussian blurring kernel with standard deviation equal to the distance between probes, such a strategy could achieve a percent score that is nearly identical to that of the human observers (see Table 1). However, no trial–by–trial correlation was found between this strategy and the observers' performance. Thus, we regard it as unlikely observers were using such a strategy.

5.2 Lighting models

A key question raised by these experiments is whether observers were using a different strategy in the diffuse lighting condition than in the point source conditions. Following the experiments, we debriefed observers about the purposes of the experiment. Many observers spontaneously claimed that they had been unaware of the lighting conditions changing between trials, and expressed surprise at this manipulation. Some observers reported that certain images had lower contrast than others – the diffuse source images do have lower contrast than the point source images – but they were unable to report any other differences. Many observers also reported that the response times were so short that they were only able to attend to a local neighborhood of each probe point. If they had more time, they argued, they would have noticed how some images had a sharp shadow (the point source condition) and other did not (the diffuse lighting condition). It remains possible that the lighting condition was identified subconsciously, and that observers indeed used a different strategy from one lighting condition to the next. This issue will need to be addressed in future experiments.

Acknowledgments

The authors thank I. Bülthoff, V. Blanz, A. Blake, V. Franz, Z. Liu, J. Stewart and N. Troje for assistance. This research was supported by the DAAD and by an Alexander von Humboldt Research Fellowship to M. Langer.

References

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

Berbaum, K., Bever, T., & Chung, C. (1984). Extending the perception of shape from known to unknown shading. *Perception*, **13**, 479–488.

Brewster, D. (1826). On the optical illusion of the conversion of cameos into intaglios and of intaglios into cameos, with an account of other analogous phenomena. *Edinburgh Journal of Science*, **4**, 99–108.

Bülthoff, H., & Mallot, H. (1988). Interaction of Depth Modules: Stereo and Shading. *Journal of the Optical Society of America*, **5**, 1749–1758.

Christou, C. G., & Koenderink, J. J. (1997). Light Source Dependence in Shape from Shading. *Vision Research*, **37**(11), 1441–1449.

Dosher, B. A., Sperling, G., & Wurst, S. A. (1986). Tradeoffs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure. *Vision Research*, **26**(6), 973–990.

Dupuis, P., & Oliensis, J. (1994). An Optimal Control Formulation and Related Numerical Methods for a Problem in Shape Reconstruction. *Annals of Applied Probability*, **4**(2), 287–346.

E. MacCurdy, e. (1938). *The Notebooks of Leonardo da Vinci*. London: Jonathan Cape. p. 332.

Gregory, R. (1970). *The Intelligent Eye*. New York: MacGraw-Hill.

Horn, B. K. P. (1975). *The Psychology of Computer Vision*, chap. Obtaining Shape from Shading Information. New York: McGraw-Hill.

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

Howard, I. (1983). Occluding edges in apparent reversal of convexity and concavity. *Perception*, **12**, 85–86.

Howard, I., Bergstrom, S., & Ohmi, M. (1990). Shape from shading in different frames of reference. *Perception*, **19**, 523–530.

Ikeuchi, K., & Horn, B. K. P. (1981). Numerical Shape from Shading and Occluding Boundaries. *Artificial Intelligence*, **17**, 141–184.

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

Johnston, A., & Passmore, P. (1994b). Shape from shading. II: Geodesic bisection and alignment. *Perception*, **23**, 191–200.

Koenderink, J., van Doorn, A., & Kappers, A. (1992). Surface Perception in Pictures. *Perception and Psychophysics*, **52**(5), 487–496.

Koenderink, J. J. (1984). What does the occluding contour tell us about solid shape?. *Perception*, **13**, 321–330.

Koenderink, J. J., van Doorn, A. J., & Kappers, A. M. L. (1996). Pictorial surface attitude and local depth comparisons. *Perception and Psychophysics*, **58**, 163–173.

Langer, M., & Zucker, S. (1994). Shape-from-shading on a cloudy day. *Journal of the Optical Society of America A*, **11**(2), 467–478.

Larson, G. W., & Shakespeare, R. (1998). *Rendering with Radiance: the art and science of lighting visualization*. Morgan Kaufmann Publishers, Inc.

Mach, E. (1866). Über die physiologische Wirkung räumlich verteilter Lichtreize.. translated in Ratliff, F. Mach Bands: Quantitative Studies on neural networks in the retina (Holden-Day, Inc. 1965).

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

Mamassian, P., Kersten, D., & Knill, D. (1996). Categorical local-shape perception. *Perception*, **25**, 95–107.

Mingolla, E., & Todd, J. (1986). Perception of Solid Shape from Shading. *Biological Cybernetics*, **53**, 137–151.

Nicolaides, K. (1941). *The Natural Way to Draw*. Boston: Houghton Mifflin.

Norman, J. F., & Todd, J. T. (1996). The discriminability of local surface structure. *Perception*, **25**, 381–398.

Peli, E. (1990). Contrast in complex images. *Journal of the Optical Society of America A*, **7**(10), 2032–2040.

Pentland, A. P. (1982). Finding the Illuminant Direction. *Journal of the Optical Society of America*, **72**, 448–455.

Pentland, A. P. (1990). Linear shape from shading. *International Journal of Computer Vision*, **4**, 153–162.

Ramachandran, V. (1988). Perception of Shape from Shading. *Nature*, **331**(6152), 163–165.

Reichel, F. D., Todd, J. T., & Yilmaz, E. (1995). Visual discrimination of local surface depth and orientation. *Perception and Psychophysics*.

Reichel, F. R., & Todd, J. T. (1990). Perceived Depth Inversion of Smoothly Curved Surfaces Due to Image Orientation. *Journal of Experimental Psychology: Human Perception and Performance*, **16**(3), 653–664.

Rittenhouse, D. (1786). Explanation of an optical deception. *Transactions of the American Philosophical Society*, **2**, 37–43.

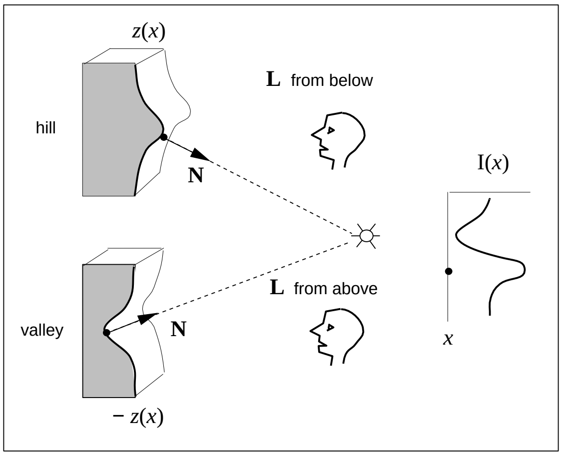
Stewart, A. J., & Langer, M. S. (1997). Towards Accurate Recovery of Shape from Shading Under Diffuse Lighting. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **19**, 1020–1025.

Sun, J., & Perona, P. (1998). Where is the sun?. *Nature Neuroscience*, **1**(3), 183–184.

Todd, J., & Mingolla, E. (1983). Perception of Surface Curvature and Direction of Illumination from Patterns of Shading. *Journal of Experimental Psychology: Human Perception and Performance*, **9**(4), 583–595.

Todd, J., & Reichel, F. (1989). Ordinal Structure in the Visual Perception and Cognition of Smoothly Curved Surface. *Psychological Review*, **96**(4), 643–657.

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



(1a)