Are discrimination thresholds a valid measure of variance for judgments of slant from texture?

James T. Todd Department of Psychology, Ohio State University, Columbus, OH, USA

James T. Christensen Air Force Research Laboratory, Dayton, OH, USA

Kevin C. Guckes Department of Psychology, Ohio State University, Columbus, OH, USA

There have been many experiments reported in the literature that have employed discrimination procedures to estimate the variance of observers' slant judgments from texture and binocular disparity, both individually and in combination. The research described in the present article identifies two serious methodological flaws in these studies. Although discrimination thresholds can be influenced by the variance of observers' slant estimates, they can also be affected by systematic biases in observers' judgments, and the presence of 2D cues that are irrelevant to the perception of slant. A series of five experiments is reported to show that: (1) the slants of surfaces specified by texture gradients can be systematically underestimated; (2) surfaces specified by texture gradients appear significantly less slanted than equivalent surfaces specified by binocular disparity; (3) the difference in bias between observers' slant judgments from stereo and texture may be more important than their relative variance in determining how these cues are weighted when presented in combination; (4) observers may be less sensitive to variations in apparent slant from texture than they are to variations in 2D cues that are unrelated to the perception of slant; and (5) these 2D cues may be the primary source of information for discriminating images of textured surfaces. These findings provide strong evidence that the results of prior discrimination studies have been misinterpreted because of the confounding effects of bias and/or 2D cues, and that this has resulted in several questionable conclusions that have been broadly accepted within the field.

Keywords: shape, depth, slant, texture, cue integration

Citation: Todd, J. T., Christensen, J. T., & Guckes, K. C. (2010). Are discrimination thresholds a valid measure of variance for judgments of slant from texture?. *Journal of Vision, 10*(2):20, 1–18, http://journalofvision.org/10/2/20/, doi:10.1167/10.2.20.

Introduction

One of the more active areas of research in 3D vision during the past decade concerns how different sources of information (e.g., texture, shading, motion, or disparity) are combined to produce a single unitary percept of an observed object in space. Many researchers have argued that the most sensible way of combining cues is a linear weighted average, in which the weights are determined by the relative reliability of each cue (e.g., Ernst & Banks, 2002; Hillis, Ernst, Banks, & Landy, 2002; Hillis, Watt, Landy, & Banks, 2004; Knill & Saunders, 2003). For example, consider the analysis of surface slant from texture and binocular disparity. If errors in observers' slant estimates from these cues are uncorrelated and have variances (σ_t^2 and σ_d^2) that are inversely related to the respective cue reliabilities, then this strategy will ensure that the variance of slant estimates when both cues are combined (σ_c^2) will be smaller than the variances obtained when either cue is presented individually, as described by

$$\sigma_c^2 = \frac{\sigma_d^2 \sigma_t^2}{\sigma_d^2 + \sigma_t^2}. \quad (1)$$

In order to test this hypothesis empirically, it is necessary "to acquire valid measures of the reliabilities of single-cue *slant* estimates" (Hillis et al., 2004). One popular way of doing this is to measure slant discrimination thresholds, "based on the assumption that thresholds are proportional to the standard deviation of internal slant estimates" (Knill & Saunders, 2003). In the present article, we will provide theoretical arguments and empirical evidence that slant discrimination thresholds are an invalid measure of variance because they are confounded by other factors that can influence performance, and that the use of this procedure has led to several questionable conclusions that have gained wide acceptance in the field. Our discussion will be primarily concerned with the analysis of slant from texture, which is a frequently used condition in experiments on cue combination, but the problems we will reveal can also be generalized to experiments involving other sources of information.

The perception of surface slant from texture

Let us begin by summarizing some prior research on the perception of shape from texture. There are several different properties of optical texture patterns that have been identified as possible sources of information for the perception of local slant. One popular approach for estimating slant is based on an assumption that variations in reflectance on a surface are statistically isotropic. In the special case of polka dot textures, as shown in Figure 1, the optical slant (τ) at the center of each element can be determined by the following equation:

$$\cos(\tau) = \frac{\omega}{\lambda}, \tag{2}$$

where λ and ω are the major and minor axes of its optical projection. This is often referred to as the foreshortening cue. Similar computations can also be performed for less regular isotropic textures from the distribution of edge orientations in each local image region (Aloimonos, 1988; Blake & Marinos, 1990; Blostein & Ahuja, 1989; Marinos & Blake, 1990; Witkin, 1981), or from the relative anisotropy of their local amplitude spectra (Bajcsy & Lieberman, 1976; Brown & Shvayster, 1990; Krumm & Shafer, 1992; Sakai & Finkel, 1995; Super & Bovik, 1995). One important weakness of this cue relative to others is that it cannot reveal the sign of slant—only its magnitude.

An alternative approach that does not share this weakness is to estimate surface slant by measuring the changes of optical texture across different local neighborhoods of an image, based on an assumption that the texture on a physical surface is statistically homogeneous. As was first demonstrated by Purdy (1958), the optical slant (τ) in a given local region can be determined by the following equation:

$$\tan(\tau) = \frac{2(\lambda_1 - \lambda_2)}{\delta(\lambda_1 + \lambda_2)}, \tag{3}$$

where δ is the projected distance between neighboring optical texture elements in the direction that slant is being estimated, and λ_1 and λ_2 are the projected lengths of those texture elements in a perpendicular direction (see Figure 1). In the limit of an infinitesimally small δ, the right side of Equation 2 is equal to the normalized depth gradient (Purdy, 1958, Equation 14; Gårding, 1992, Equation 33). This is sometimes referred to as the scaling cue. Similar computations can also be performed on less regular textures from the affine correlations between the amplitude spectra in neighboring image regions (Clerc & Mallat, 2002; Malik & Rosenholtz, 1994, 1997) or from systematic changes in the distributions of edges (Gårding, 1992, 1993).

An extensive series of experiments and simulations was performed by Knill (1998a, 1998b) in an effort to determine the relative importance of these different possible texture cues for slant discrimination judgments. For example, one technique he employed involved manipulating the relative reliability of the cues by adding random variations to some local texture properties but not others. From the results of these studies, Knill concluded that observers' slant estimates are based primarily on the foreshortening cue. He was also the first to discover that discrimination thresholds for shallow slants are an order of magnitude larger than those obtained for steep slants. This finding has proven to be important for subsequent research on cue integration because it suggests that texture should be weighted more heavily for steep slants than for shallow slants in relation to other cues.

A more recent series of studies by Todd, Thaler, and Dijkstra (2005) and Todd, Thaler, Dijkstra, Koenderink, and Kappers (2007) has produced a contradictory pattern of results. They used adjustment tasks in which observers were asked to duplicate the apparent variations in depth on a surface. The results revealed that the ability to distinguish slants from texture requires relatively large viewing angles, which provides strong evidence that observers' perceptions cannot be based on computational analyses within small local neighborhoods. In light of this finding, Todd et al. (2007) proposed a new source of information, called scaling contrast, which is defined by the following equation:

$$\text{Scaling contrast} = \frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}}, \tag{4}$$

where λ_{max} and λ_{min} are the lengths of the largest and smallest texture elements over the entire extent of a visible surface. This measure is similar to Equation 3, but

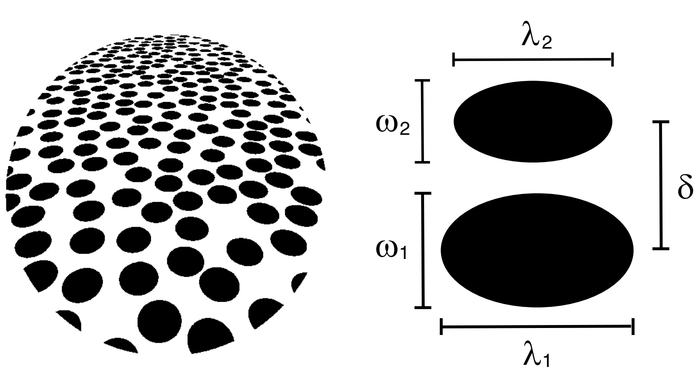

Figure 1. An image of a slanted plane that is perceptually specified by optical texture, and some local measures of texture that are useful for the computation of surface slant.

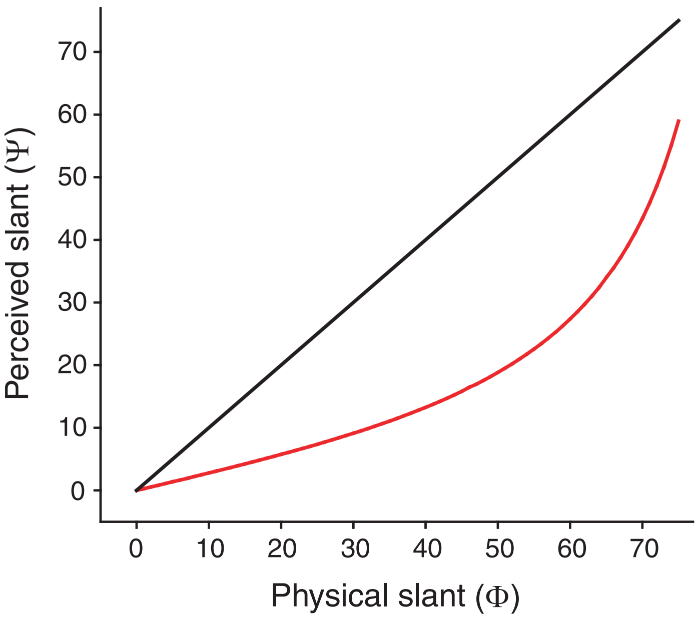

Figure 2. Some hypothetical functions by which perceived slant may be related to the ground truth. The black curve shows the expected pattern of apparent slants for an unbiased observer. The red curve shows the pattern that would be expected based on the scaling contrast model proposed by Todd et al. (2007).

it is designed to evaluate the variations in scaling over large regions of visual space, rather than small local neighborhoods. Todd et al. (2007) found that it is highly correlated with observers' shape judgments over a wide range of conditions.[1] Another interesting finding from these studies is that the variance in observers' settings did not change dramatically with slant, as has been reported by Knill (1998a, 1998b) and others for slant discrimination thresholds. If anything, the changes in variance as a function of slant were in the opposite direction.

The problem of bias

To better understand how discrimination and adjustment tasks could produce such different results, it is useful to consider the psychometric function by which perceived slant (Ψ) varies with physical slant (Φ). The red curve in Figure 2 shows the pattern of perceived slant predicted by scaling contrast for a planar surface patch within a 15° field of view. Other things being equal, the slant discrimination threshold (T) for any given base slant is determined by the following equation:

$$T \propto \sigma / \frac{d\psi}{d\phi}, \tag{5}$$

where σ is the standard deviation of observers' slant judgments, and $d\psi/d\phi$ is the slope of the psychometric function at that base slant. If the function is curved, as predicted by scaling contrast, then the slant thresholds may vary as a function of the base slant because of the changing slope, even if the variance of observers' judgments remains constant across all slants.

How then does one determine in any given experiment whether threshold differences at different base slants are due to differences in the variance of observers' judgments, or to different slopes in the function that relates physical and perceived slants. The way this has been dealt with in the previous literature is to "assume an unbiased observer". That is to say, it is assumed that the psychometric function is linear with a slope of one, or, stated differently, that the observers' judgments are on average veridical. Unfortunately, there has been little or no discussion of whether that assumption is reasonable, or what the consequences would be for one's empirical conclusions if it turned out to be incorrect.

Figure 3 shows some of the stimuli used by Knill and Saunders (2003) in a cue combination study involving texture and stereo, which should help clarify these issues. Note how the slants of the 0° and 30° surfaces are difficult to distinguish. One possible explanation of this based on variance is that an observer's slant estimates change randomly from moment to moment over a relatively large range, such that at any given instant there is a substantial probability that the 0° surface will appear more slanted than the 30° surface. The alternative explanation based on bias is that the slant of the 30° surface is perceptually underestimated so that it

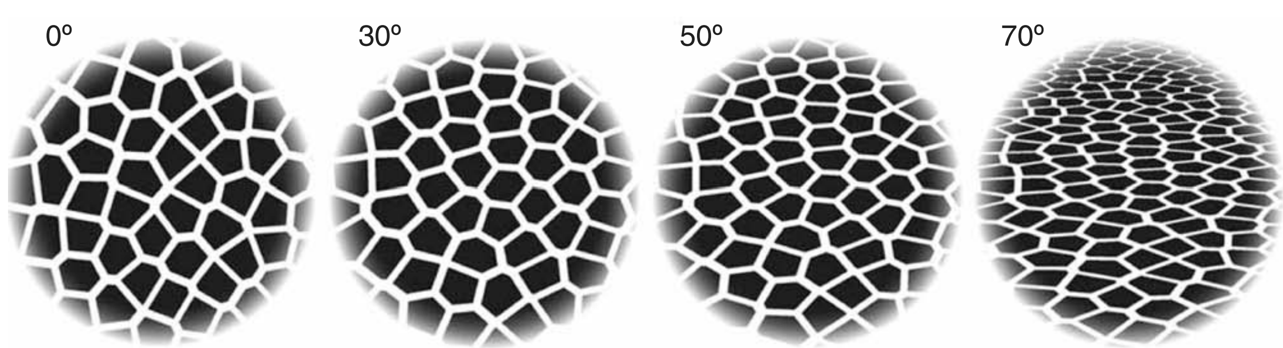

Figure 3. Some stimuli used by Knill and Saunders (2003) in a slant discrimination study.

appears reliably to be very close to frontoparallel. Readers are invited to judge for themselves which of these explanations is more consistent with their own perceptual experiences for these displays. Another important thing to note in Figure 3 is that the 30° difference between the surfaces with slants of 0° and 30° appears perceptually smaller than the 20° difference between the surfaces with slants of 50° and 70°, which is the opposite of what one would expect if the perception of these surfaces were unbiased. This type of reversal can only occur when there is a nonlinear relationship between physical and perceived slants, such as the one depicted in Figure 2.

The research described in the present article was designed to provide more rigorous empirical support for these anecdotal observations using several different methodological procedures. The results of these experiments will demonstrate that the apparent slants of planar surfaces are based primarily on scaling contrast, and that the heights and aspect ratios of optical texture elements have little or no influence on perceived slant. These results will also confirm that the apparent slants of planar surfaces from texture are systematically underestimated for relatively small fields of view (as have been used in most slant discrimination studies), and that they vary as a curvilinear function of the depicted physical slant. Other experiments will show that surfaces specified by texture gradients appear significantly less slanted than equivalent surfaces specified by binocular disparity. This differential bias between stereo and texture has important theoretical implications. If the goal of an optimal cue integration strategy is to minimize perceptual errors, then the best way to achieve that for texture and stereo would be to assign little or no weight to the texture information. Our empirical results will confirm that prediction. Finally, these experiments will also show that observers may be less sensitive to variations in apparent slant from texture than they are to variations in the heights or aspect ratios of optical texture elements, which are unrelated to the perception of slant, and that these 2D cues may be the primary source of information for discriminating images of textured planar surfaces.

Experiment 1

A fundamental prediction of the scaling contrast model is that the magnitude of perceived depth or slant should vary systematically as a function of the viewing angle with which a surface is observed, or, in the case of pictorial images, it should vary as a function of the camera angle with which the image was rendered (see Figure 4). Figure 5 shows 11 images of planar surfaces with slants of 30°, 50°, or 70° that were rendered in four different ways: Under orthographic projection, and with camera angles of 15°, 30°, or 60°. Note in particular that most of these surfaces appear to have slants that are substantially less than the ground truth (even when viewed monocularly), and that the magnitude of apparent slant increases with camera angle, as predicted by the model. It is also important to highlight in this regard that a veridical estimate of slant could be obtained for all of these images by exploiting the foreshortening cue (see Equation 2) for texture elements in the center of each display. Experiment 1 was designed to test the predictions of the scaling contrast and foreshortening models using an adjustment task, in which observers could record the magnitude and direction of perceived slant for images of textured planar surfaces with varying slants that were rendered with different camera angles like those in Figure 5.

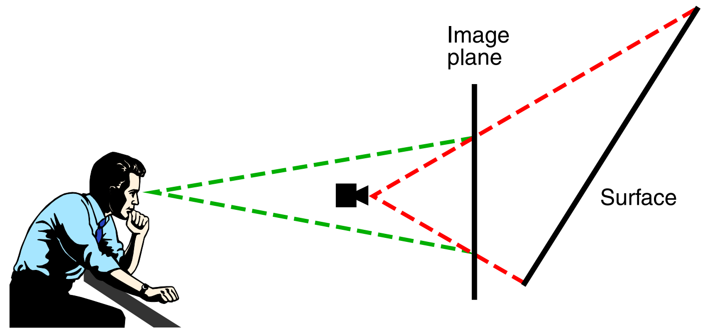

Figure 4. A schematic diagram of an image of a surface that is viewed with a different visual angle from the one at which it was photographed or rendered. The camera angle in this example is represented by dashed red lines, and the viewing angle is represented by dashed green lines.

Methods

Subjects

Seven observers participated in the experiment, including the authors (JC, KG, and JT) and four others (DL, AP, LT, and DW) who were naive about the issues being investigated. They all had normal or corrected-to-normal visual acuity.

Apparatus

The experiment was conducted using a Dell Dimension 8300 PC with an ATI Radeon 9800 PRO graphics card. Images were presented on a 21-in. CRT with a spatial resolution of 1280 × 1024 pixels. They were viewed monocularly with an eye patch at a distance of 112 cm, and a chin rest was used to constrain head movements. This experimental setup was also used in Experiments 2–5.

Stimuli

Images of planar surface patches were rendered using 3D Studio Max by Kinetix with procedural rounded tile

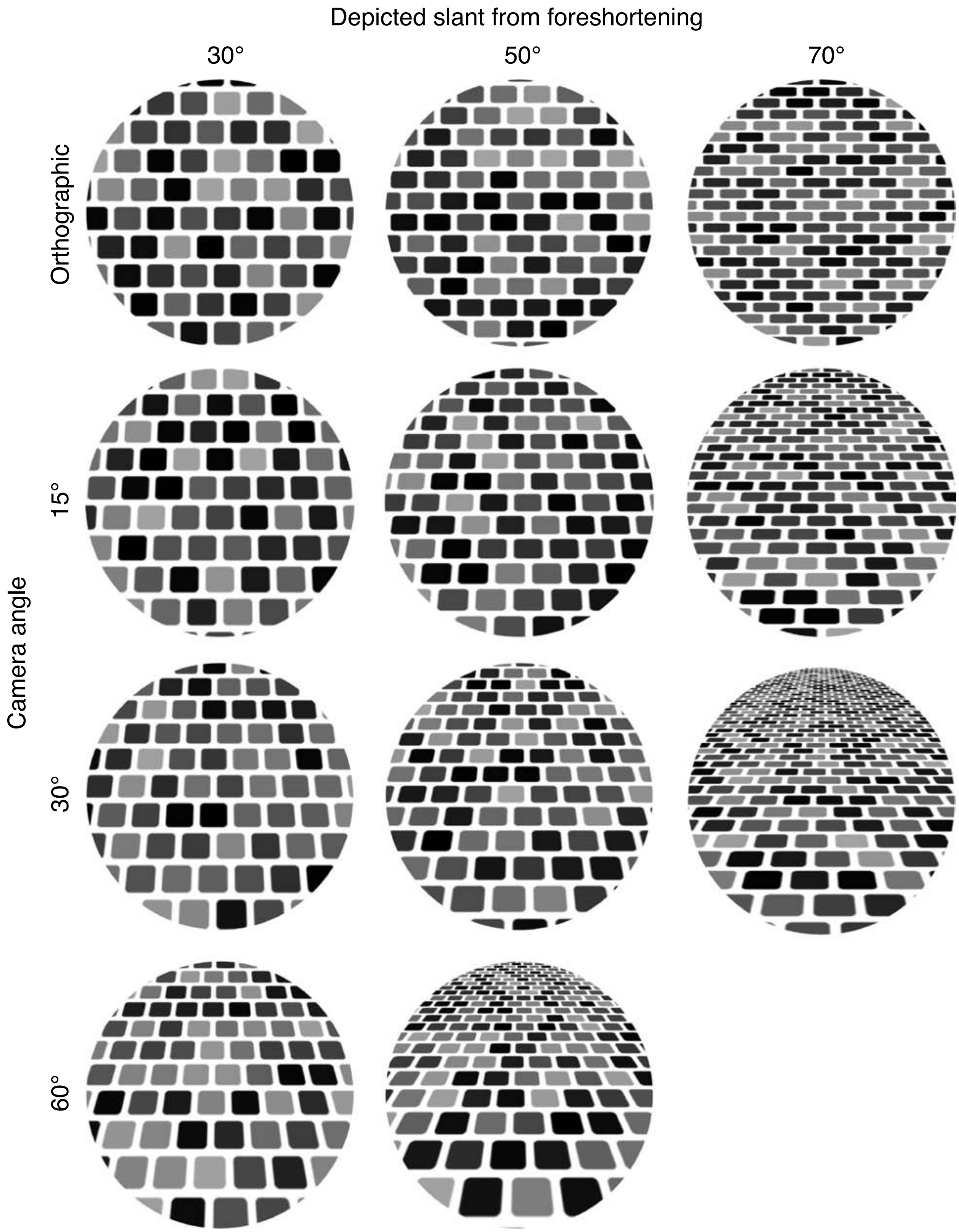

Figure 5. Some sample stimuli from the present series of experiments. The slants of the depicted surfaces were systematically varied, and the images were rendered with a variety of different camera angles. The appearance of slant in these images is enhanced when they are viewed monocularly.

Figure 6. The response screen for Experiments 1 and 3. Observers adjusted the orientation of a line so that it matched the apparent slant of a textured surface presented in a different screen.

textures that were created with the DarkTree 2.5 texture plug-in by Darktree Studios. The images were all clipped so that they would only be visible within a circular aperture in the image plane with a diameter of 15°. The color of each tile was selected randomly on every trial over a range from black to medium gray, and the tiles were surrounded by white grout. The scales of the textures in different conditions were all normalized so that eight tiles would always be visible along a horizontal cross-section through the center of each aperture. The depicted surfaces could have eight possible slants of ±0°, ±30°, ±50°, or ±70° relative to the frontoparallel plane, and they could be rendered with four possible camera angles including orthographic projection, 15°, 30°, or 60° (see Figure 5). Because it is not mathematically possible to have a 60° view of a surface with a 70° slant, that combination of parameters was excluded from the design. In addition, because images of surfaces with 0° slants are the same for all camera angles once the scale of the texture is normalized, we did not vary the camera angle for the 0° slant conditions. Taking these exclusions into account, the stimulus set consisted of 24 possible combinations of camera angle and slant.

Procedure

Each trial began with an image of a textured surface presented within the 15° circular aperture. After viewing this image, observers could perform a mouse click to replace it with one that depicted a human head in profile gazing at a line, which was intended to represent how that surface would appear if viewed from the side (see Figure 6). Observers were required to adjust the orientation of this line using the mouse so that it matched the apparent slant of the surface in the first screen (Van Ee & Erkelens, 1996). They were allowed to toggle back and forth as many times as necessary until they were satisfied with their adjustments. All observers agreed that this task was quite natural and that they had a high degree of confidence in their settings. Each observer made five adjustments for each of the 24 possible combinations of slant and camera angle over two experimental sessions.

Results and discussion

A preliminary assessment of the data revealed that there were negligible differences between the judgments obtained for positive and negative slants, so we collapsed these conditions by using their absolute values for all subsequent analyses. Figure 7 shows the average settings over all six observers for each combination of depicted slant magnitude and camera angle. This pattern of results is also representative of those produced by each of the individual observers. The 15° camera condition has been highlighted in black, because that is the one that matched the actual viewing angle for these displays, and it also matches the viewing geometry employed by Knill and Saunders (2003). Note that the depicted slants of these 15° displays were severely underestimated, and that there is a noticeable curvature in the psychometric function. These results are obviously incompatible with an assumption that observers' judgments are unbiased (see also Todd et al., 2005, 2007). Because the textures employed by Knill and Saunders (see Figure 3) were less regular than the rounded tile textures used in this study, it is quite likely that the underestimation of slant for their stimuli would be even greater.

Another important conclusion from these data is that the texture foreshortening cue may have little or no impact on observers' perceptions of slant (see also Gillam, 1970).

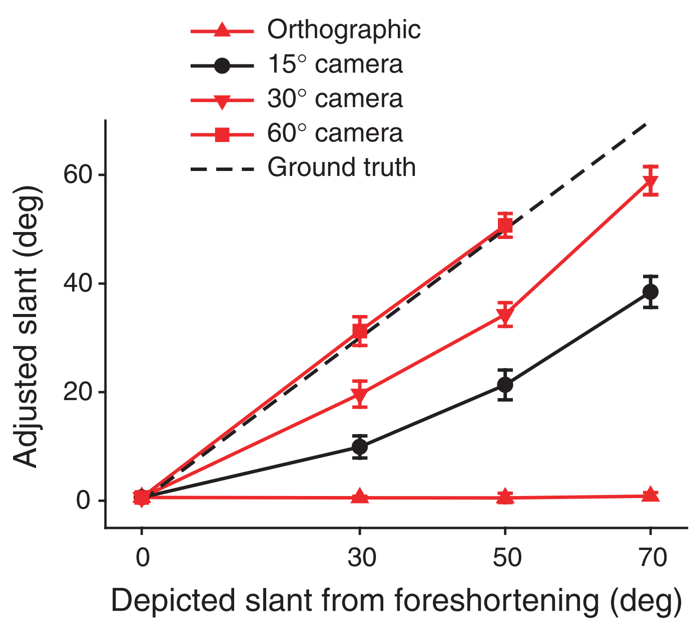

Figure 7. Average judged slants from Experiment 1 as a function of the depicted slant for all of the different camera angles.

Although this cue was available to obtain veridical slant estimates for the displays rendered under orthographic projection, the observers judged all of these displays as frontoparallel surfaces, regardless of the depicted slant. This conclusion is also supported by other recent findings on the perception of 3D shape from anisotropic textures (Todd, Oomes, Koenderink, & Kappers, 2004).

Other possible sources of information for estimating the slants of these stimuli include local scaling gradients, as described in Equation 3, or the contour convergence cue described by Braunstein and Payne (1969) and Backus and Saunders (2006). If observers' slant judgments had been based on either of these analyses, the orthographic displays would have all appeared frontoparallel (which they did), the stimuli rendered with 15° camera angles would have been perceived veridically, and the slants of those with 30° or 60° camera angles would have been severely overestimated (i.e., for surface slants of 30°, 50°, and 70° with a 15° field of view, the gradient cue defined by Equation 3 would specify slants of 51°, 69°, and 80°, respectively, for displays rendered with a 30° camera angle, and 75°, 83°, and 87° for displays rendered with a 60° camera angle). Note that none of these predictions (except for the orthographic ones) are confirmed by the data. These analyses are also incapable of explaining the curvilinear relationship between physical and perceived slants for the displays with 15° camera angles. We will examine this particular aspect of the results in more detail in Experiment 4.

The results are consistent, however, with the predictions of the scaling contrast model, as shown in Figure 8. An analysis of linear regression revealed that this model accounts for over 94% of the variance in the observers' slant judgments among the different combinations of camera angle and depicted slant. Nevertheless, upon closer examination of this figure, the residuals do not appear to be random. Todd et al. (2007) found that the scaling contrast model overestimates the magnitude of perceived slant for displays that are too highly foreshortened in peripheral regions, presumably because the spatial frequency in those regions is too high to effectively compute the relative lengths of individual texture elements (e.g., see the 70° surface that was rendered with a 30° camera in Figure 5). Thus, it is likely that the effective scaling contrast was reduced for the displays in the lower right corner of Figure 5 that had the highest frequency texture variations in their uppermost regions.

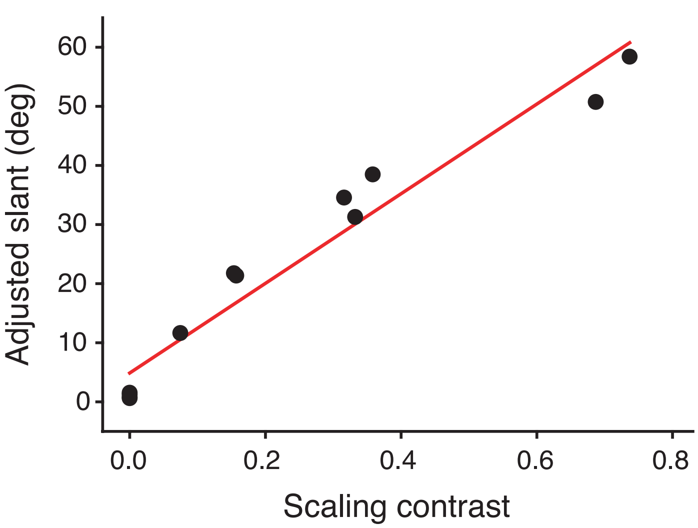

Figure 8. The average judged slants from Experiment 1 as a function of scaling contrast.

Let us now consider the pattern of variance in observers' judgments. One of the most well-accepted conclusions in the perception of slant from texture is that observers' judgments of steep slants are less variable than their judgments of shallow slants. The empirical support for this conclusion comes from slant discrimination studies (e.g., Hillis et al., 2004; Knill, 1998a, 1998b; Knill & Saunders, 2003), but the thresholds obtained with that procedure do not provide a pure measure of variance, because they are confounded with the effects of bias. The adjustment task in the present experiment allows these factors to be easily disentangled. The mean of the observers' settings allows a pure estimate of their systematic biases, whereas the standard deviation of their settings provides a pure estimate of variance.

Let us first consider the results for the 0° and orthographic conditions in which the surfaces were perceived as frontoparallel. All of the observers reported during their debriefing sessions that these were the easiest conditions to judge, and their objective response data confirm that quite clearly. Over all observers and simulated slants there were 280 different responses for the orthographic displays, and the standard deviation of those responses was only 2.1°. In other words, when observers' slant judgments were measured using an adjustment task in the present study, there was negligible variance among the different observers and different simulated slants for the displays that appeared to have a frontoparallel orientation. For the perspective displays that appeared slanted in depth, the average standard deviation of observers' settings in each condition ranged from 5° to 7°, and there was no significant correlation between the variance of their settings and the magnitude of perceived slant. In order to compare the variances in the orthographic and perspective conditions, we first performed a fourth root transformation on the standard deviations of each observer in each condition to normalize the distributions (Hawkins & Wixley, 1986), and then performed an ANOVA on the transformed data. The results revealed that the displays that appeared frontoparallel produced significantly smaller variances than the ones that appeared slanted in depth, $F(1,6) = 67.99$, $p < 0.001$.

These results stand in stark contrast to the conventional wisdom that the variance of slant estimates for frontoparallel surfaces is an order of magnitude higher than the variance obtained for judgments of surfaces with 70° slants. Our findings indicate that the variance of slant judgments for frontoparallel surfaces is reliably smaller than the variance obtained for surfaces that appear slanted in depth (see also Todd et al., 2005). Indeed, the estimated variance for frontoparallel surfaces using our adjustment task was approximately 20 times smaller than what has been reported previously based on discrimination measures. A likely explanation for these strikingly conflicting results is that the discrimination data on which the conventional wisdom is based has been misinterpreted by attributing threshold variations to differences in variance, rather than to differences in bias.

It is interesting to note when considering these findings that a similar response task has been employed previously by Watt, Akeley, Ernst, and Banks (2005) to investigate the perception of surface slant from texture and binocular disparity. Their displays had a 35° field of view, which should have attenuated the underestimation of slant from texture relative to the 15° displays used by Knill and Saunders (2003) and in the present study. Nevertheless, two of their three observers underestimated the slants of the textured surfaces by 35% and 75%, respectively. Despite these results, Watt et al. were unwilling to abandon their assumption that "perceived slant was veridical", and attributed these errors instead to the "unknown mapping between perceived slant and the response settings". In other words, they assumed that the errors were unrelated to an observer's initial perception of surface slant and must have arisen later during the process of mentally computing how the perceived surface would appear if viewed from the side. There is however a serious problem with the logic of this suggestion. The Watt et al. observer who underestimated the slants of textured surfaces by 75% produced estimates of slant from binocular disparity that were slightly overestimated. Given that the response task was exactly the same for both cues, this difference cannot be dismissed as the result of some unknown transfer function. Indeed, all three of the observers in Watt et al.'s study produced higher estimates of slant for the stereoscopic surfaces than for the monocular textured surfaces. It is simply not possible, therefore, that the slant estimates obtained for both of these cues were veridical, regardless of the unknown mapping between response settings and perceived slant.

Experiment 2

Experiment 2 was designed to explore these issues further using an alternative response task that did not require observers to estimate how a surface would appear when viewed from a different vantage point. Observers were presented with two displays that could be toggled back and forth with a mouse click: One that contained a fixed monocular image of a surface with an informative rounded tile texture, and another that contained an adjustable stereoscopic surface with an uninformative random noise texture. The task on each trial was to adjust the stereoscopic surface so that its slant appeared equal to that of a monocular textured surface.

Methods

Subjects

Four observers participated in the experiment, including one of the authors (JT) and three others (EE, DL, and AP) who were naive about the issues being investigated. They all had normal or corrected-to-normal visual acuity and stereo acuity.

Apparatus

The basic setup was the same as described for Experiment 1. Dichoptic presentation of the left- and right-eye images was achieved using CrystalEyes liquid crystal shutter glasses. The monitor refresh rate was 120 Hz, so that each eye's image was redrawn at 60 Hz. To minimize cross-talk between left- and right-eye images, all of the stimuli were rendered in varying shades of red, because the red phosphor of the CRT had the fastest decay rate. The displays were viewed at a distance of 112 cm, and a chin rest was used to constrain head movements.

Stimuli

Planar surface patches of varying slant were simulated using 3D Studio Max by Kinetix. For the monocular displays, we used the same procedural rounded tile textures as in Experiment 1. These surfaces were rendered using the XidMary plug-in by Habware to create stereograms that were appropriate for the CrystalEyes glasses, such that the same exact image would be presented to both eyes. The images were all clipped so that they would only be visible within a circular aperture in the image plane with a diameter of 15°. The color of each tile was selected randomly on every trial over a range from bright to medium red, and the tiles were surrounded by black grout. As in Experiment 1, the scales of the textures in different conditions were all normalized so that eight tiles would always be visible along a horizontal cross-section through the center of each aperture. The depicted surfaces could have eight possible slants of ±10°, ±30°, ±50°, or ±70° relative to the frontoparallel plane, and they could be rendered with three possible camera angles of 15°, 30°, or 60° (see Figure 5). The 70° slant was only used with the 15° camera angle.

For the stereo displays, we used a thresholded Gaussian noise texture that was always projected onto a surface in a direction that was parallel to the cyclopean line of sight. This ensured that there would never be any foreshortening in the rendered images regardless of the depicted slant. The scale of the texture was adjusted so that it was large enough to have a perceptible mesostructure (Koenderink & Van Doorn, 1996), and small enough to make it difficult for the texture elements to be perceptually individuated. The surfaces were all clipped by their intersection with a cone, whose apex was at the same position as the camera, and whose axis was coincident with the line of sight. This creates a circular boundary in the cyclopean view, but the boundary is distorted as a function of slant in the left- and right-eye images of the stereogram. Had we used a fixed circular boundary with zero disparity, it would have produced inhibition between the boundary and nearby texture elements on the surface. All of the stereoscopic stimuli were rendered under orthographic projection using two cameras with a vergence angle of 3.22°, which matches the vergence angle in our experimental setup when fixating on the center of the display screen.[2] They included 151 images whose depicted slants varied in 1° increments from −75° to 75°. During the actual experiment, these could be swapped in and out of the graphics frame buffer in real time under the control of a handheld mouse. All of these stereoscopic stimuli had different random textures to eliminate any information from first-order optical motion during the adjustment process. This resulted in the perceptual impression of a moving slanted surface with a scintillating dynamic noise texture.

Procedure

Each trial began with the presentation of a rounded tile stimulus, and observers were instructed to close one eye to evaluate its apparent slant. After viewing this image, observers could perform a mouse click to replace it with a random noise stereogram, which they were required to view binocularly. With a small amount of practice, all of the observers learned to efficiently coordinate their mouse clicks with the opening and closing of their non-dominant eye. Observers were required to adjust the slant of the stereoscopic stimulus using the mouse so that it appeared to have an identical slant as the monocularly viewed surface. They were allowed to toggle back and forth as many times as necessary until they were satisfied with their adjustments. All observers agreed that this task was straightforward and that they were confident in their settings. To summarize the overall experimental design, there were 20 possible conditions in which the monocular displays were presented with different combinations of slant and camera angle. Each observer made five adjustments for each condition over two experimental sessions.

Results and discussion

As in Experiment 1, there were no significant differences between the judgments obtained for positive and negative slants, so we collapsed these conditions by using their absolute values for all subsequent analyses. The average standard deviation of observers' settings in each condition ranged from 5° to 9°, and there were no significant correlations between the variance of their settings and the magnitude of perceived slant. This confirms the results of Experiment 1 and those reported earlier by Todd et al. (2005), and it provides strong evidence that previous conclusions about how reliability varies with slant may need to be reexamined.

Figure 9 shows the average settings over all four observers for each combination of depicted slant magnitude and viewing angle. This pattern of results is also representative of those produced by each of the individual observers. The 15° camera condition has again been highlighted in black, because that is the one that matched the actual viewing angle for these displays. As in Experiment 1, these 15° displays were severely underestimated relative to the stereo matching stimuli, and there is a noticeable degree of curvature in the psychometric function by which the two cues are related. These data show clearly that perceived slant from texture for any given surface may be quite different from the perception of slant that is produced when the same surface is specified by binocular disparity. For example, in the 50° slant condition with a 15° camera angle, the magnitude of apparent slant from texture and stereo differed by a factor of two. That is to say, when compared to a textured surface with a 50° slant, the slant of the stereoscopic adjustment stimulus needed to be reduced to 26° in order to appear perceptually equivalent. It is also interesting to note, however, that the relative gain between these cues can be reversed by increasing the camera angle to 60° for the monocular stimuli.

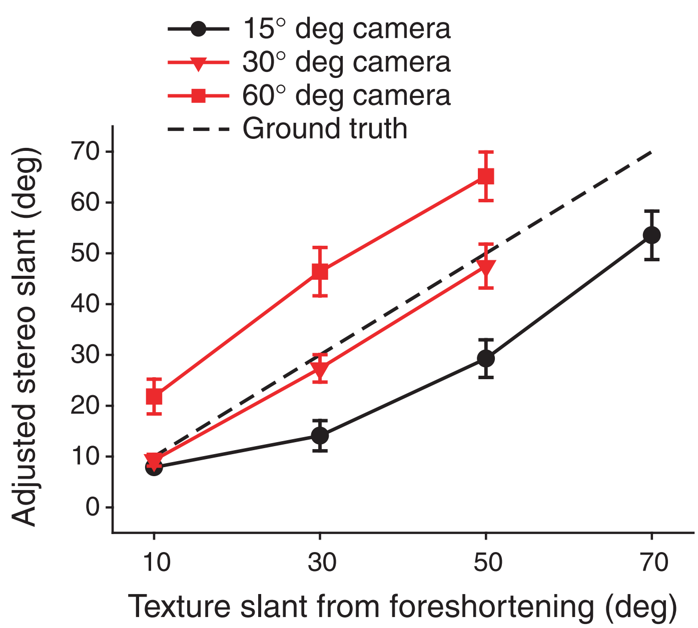

Figure 9. The average judged slants from stereo in Experiment 2 that appeared equal to the apparent slant of a textured surface that was viewed monocularly.

There have been several previous studies that have used a direct surface matching procedure like the one reported here in order to compare the apparent shapes of objects defined by different combinations of cues (Domini & Caudek, 2010; Tittle, Norman, Perotti, & Phillips, 1998; Tittle & Perotti, 1997; Todd & Perotti, 1999). There are also many others that have compared objects indirectly using other types of adjustment tasks (e.g., Bradshaw, Parton, & Glennerster, 2000; Tittle & Braunstein, 1993; Tittle, Todd, Perotti, & Norman, 1995; Todd & Norman, 2003; Watt et al., 2005). Almost all of these studies have obtained the same basic pattern of results: Objects defined by different cues that have the same depicted 3D structure can appear perceptually quite different. Conversely, objects defined by different cues that have the same apparent 3D structure can have depicted physical structures that are quite different. It is important to keep in mind that these comparisons among different cue combinations have all involved procedures in which the depicted objects were judged using a single response task with the same observers in all conditions. Thus, the reported differences cannot be attributed to an unknown mapping between perceived slant and the response setting as argued by Watt et al. (2005). Rather, these results suggest that observers' perceptions of 3D structure from different combinations of cues must be differentially biased.

Experiment 3

If the goal of an optimal cue integration strategy is to minimize error in observers' estimates of 3D structure, then it makes little sense to weight cues based solely on their relative reliability, while ignoring any systematic biases that may occur. The results of Experiments 1 and 2 indicate that the judged slants of planar surfaces from texture with 15° fields of view are systematically underestimated, and that slants specified by binocular disparity can appear twice as large as the equivalent slants defined by gradients of texture. Because the differential bias for these cues is much larger than the standard deviation of observers' judgments, it follows that the optimal strategy to minimize error when both cues are combined would be to assign little or no weight to the texture information, except in cases where perceived slants from stereo are systematically overestimated. Experiment 3 was designed to explore this issue using the same adjustment task as in Experiment 1 for images of planar surfaces with varying combinations of texture and binocular disparity.

Methods

Subjects

Three observers participated in the experiment, including one of the authors (JT) and two others (DL and AP) who were naive about the issues being investigated. They all had normal or corrected-to-normal visual acuity and stereo acuity.

Stimuli

The stereoscopic stimuli were generated using the same procedures as described for Experiment 2. That is to say, they were all rendered under orthographic projection using two cameras with a vergence angle of 3.22°, and they had back-projected textures that allowed us to manipulate the depicted slant from stereo and texture independently of one another.[2] The textures employed in the stereo-texture combined conditions were the same images rendered with 15° camera angles from the monocular stimuli in Experiment 2. The depicted slants for each cue could have eight possible values of ±0°, ±30°, ±50°, or ±70° relative to the frontoparallel plane. The sign of slant was always the same for each cue, but their relative magnitudes could occur in all possible combinations. Figure 10 shows an example of a cue conflict stimulus in which the texture pattern in the cyclopean view specifies a frontoparallel surface, but the pattern of binocular disparity specifies a surface slanted in depth. We also included a stereo-only condition with the same depicted slants using the thresholded Gaussian noise texture described in Experiment 2.

Procedure

Observers judged the apparent slant of each display using the same line orientation adjustment task as described for Experiment 1. The overall experimental design included 32 possible stereo-texture combined conditions (i.e., 16 combinations of relative slant magnitude $\times$ 2 directions of slant) and eight additional texture-only conditions. Each observer made five adjustments for each condition over two experimental sessions.

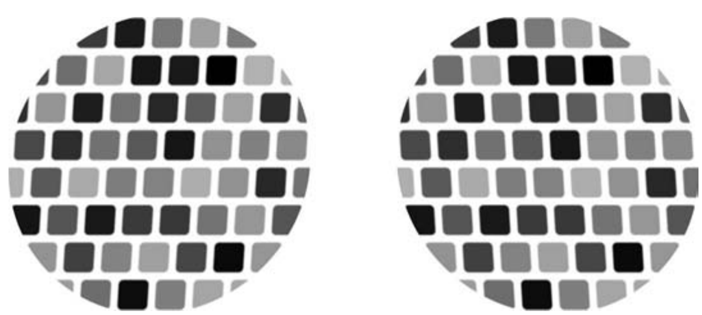

Figure 10. A cue conflict stimulus similar to the ones used in Experiment 3. The pattern of texture in the cyclopean view specifies a frontoparallel surface, whereas the pattern of binocular disparity specifies a surface slanted in depth. In the actual experiment, the tiles were colored in varying shades of red against a black background.

Results and discussion

As in Experiments 1 and 2, there were no significant differences between the judgments obtained for positive and negative slants, so we collapsed these conditions by using their absolute values for all subsequent analyses. Figure 11 shows the average settings of the observers for stereo-only displays, and the stereo-texture combined displays for which the two cues were rendered using the same ground truth. This figure also includes the average judgments of these observers from the comparable texture-only conditions in Experiment 1. It is clear from these results that the settings obtained in the stereo-only and stereo-texture combined conditions are nearly identical and also quite close to veridical for these particular viewing conditions. There is no indication from these data that the addition of texture gradients to our stimuli had any effect at all on observers' perceptions of slant.

This is in fact the optimal strategy for minimizing error given the differential biases of observers' slant estimates from these cues. This conclusion is also reinforced by the cue conflict data. Figure 12 shows the average slant settings as a function of the physical slant that was used to generate the texture patterns for each of the different stereo defined slants. Analyses of linear regression revealed that the depicted slant from stereo accounted for 97.7% of the variance in these data. The depicted slant from texture, in contrast, accounted for only 1.4% of the variance, although there does seem to be a small systematic effect for the displays with zero disparity.

There have been several previous studies that have used adjustment or magnitude estimation tasks to determine the relative weights assigned to texture and stereo when both cues are presented in combination. The one closest to the present experiment was performed by Gillam (1968) who obtained slant judgments of planar surfaces with 15° fields of view. Some of her observers behaved similarly to those in the present study in that their responses were based almost exclusively on stereo. Others produced responses that were a compromise between the two conflicting cues, though the effects of texture were still relatively small in comparison to stereo. Other investigators have examined the interaction of these cues on the apparent depth magnitudes of quadric surfaces, such as cylinders or ellipsoids (Buckley & Frisby, 1993; Johnston, Cumming, & Parker, 1993; Tittle et al., 1998). On average, the weights assigned to texture in these studies were less than 15%.

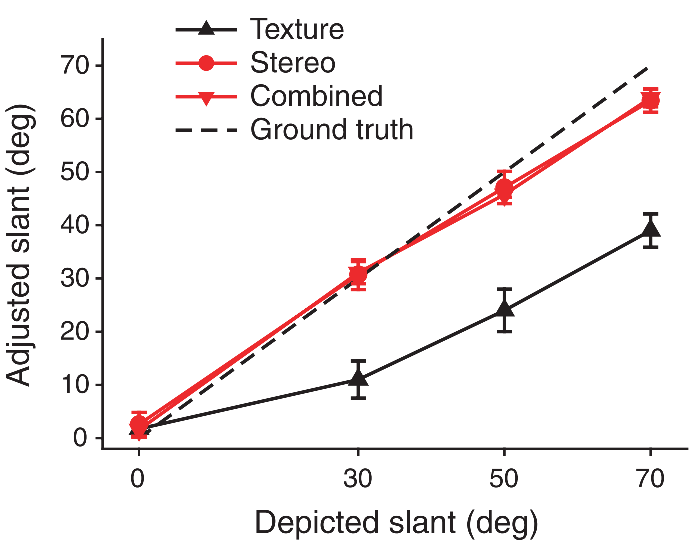

Figure 11. Average judged slants in the stereo-only and cue-consistent conditions from Experiment 3, together with the texture-only data from the same observers in Experiment 1.

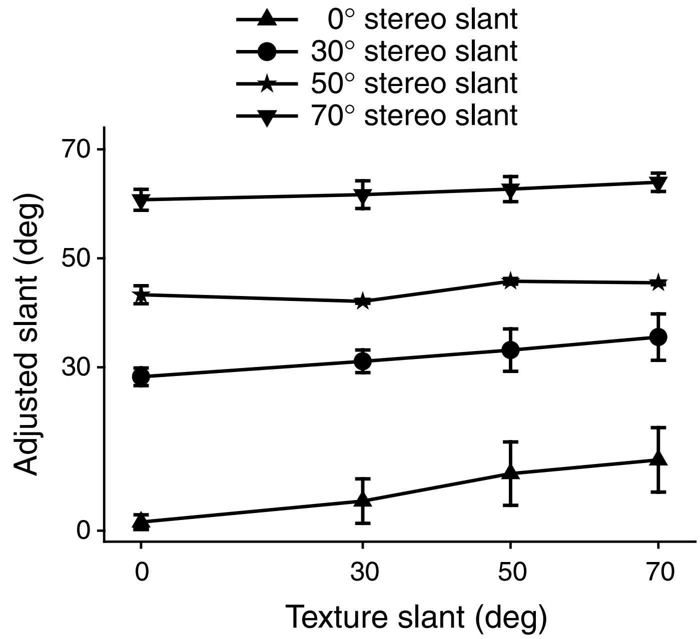

Figure 12. The average judged slants for all of the different combinations of texture- and stereo-defined slant in Experiment 3.

Experiment 4

When considering Figure 3 in the Introduction section, we highlighted a curious phenomenon that the 30° difference between surfaces with slants of 0° and 30° appears perceptually smaller than the 20° difference between surfaces with slants of 50° and 70°. This is in fact a strong prediction of the scaling contrast model for 15° fields of view, and it suggests another possible method of measuring curvature in the psychometric function relating physical and perceived slants. Suppose that an observer is asked to equate two slant intervals: A standard interval that is anchored at a base slant of 70°, and an adjustable matching interval that is anchored at some other smaller slant. According to the scaling contrast model, the magnitude of the matching interval should increase in a predictable manner as its anchor slant is decreased toward 0°. Experiment 4 was designed to test this prediction. As an additional control, we also asked observers to compare intervals in the projected heights of the tiles to see how their 2D judgments are related to their perceptions of slant.

Methods

Subjects

Six observers participated in the experiment, including the authors and three others (DL, LT, and DW) who were naive about the issues being investigated. All had normal or corrected-to-normal visual acuity and stereo acuity.

Stimuli

The displays were all generated with a 15° camera angle using the same basic procedures and textures as described for Experiment 1. We created 760 images that all had different randomizations of texture. The depicted planar surfaces had possible slants ranging from 0° to 75° in 1° increments, and 10 separate versions were created for each slant. Negative slants could be created when needed by displaying an image with a positive slant upside down.

Procedure

Each trial involved the presentation of 3 images: two that were fixed and one that was adjustable, and observers could switch from one to another using a mouse click. These displays were always presented in a fixed sequence, such that the two fixed images appeared first followed by the adjustable one. The depicted slant (τ_1) in the first image was always greater than the slant (τ_2) in the second. Observers were required to manipulate the slant (τ_3) of the adjustable stimulus using the mouse so that the apparent interval between τ_2 and τ_3 was the same as the apparent interval between τ_1 and τ_2. This was achieved by swapping images in and out of the graphics frame buffer in real time based on the mouse position. Observers were allowed to toggle through the sequence as many times as necessary until they were satisfied with their setting.

Each experimental session consisted of a series of blocks that included five repeated interval adjustments with different randomizations of texture for the two fixed images. In the first block of a session, τ_1 and τ_2 were set to 75° and 70°, respectively. The average of an observer's five settings was then used to determine the value of τ_2 in the next block and the value of τ_1 in that block was given the same value as the previous τ_2. This process was continued with progressively smaller values of τ_1 and τ_2 until the average of the five settings in a block resulted in a negative slant, at which point the session was terminated. What this produces is a set of intervals with different anchor slants that all appear subjectively equal. A second version of this task was also employed, in which observers judged intervals in the projected heights of the texture elements rather than slant. Each of the six observers performed three sessions for each of the two response tasks.

Results and discussion

It is important to keep in mind that the number of blocks in a session and the standards employed after the first block were all dependent on the observers' responses. On average, the observers required approximately six blocks to complete a session for the slant adjustment task, and approximately nine blocks for the tile height adjustment task. To smooth out some of the noise in the data, the observers' judgments were separated into six bins to combine all the trials with anchor slants (τ_2) between 0° and 20°, 21° and 30°, 31° and 40°, 41° and 50°, 51° and 60°, and 61° and 70°. Note that the lower bin is larger because there were relatively few blocks with anchor slants less than 21°. The black curves in Figure 13 show the average adjusted intervals in each bin over all observers and sessions for each of the two tasks.

For the slant adjustment task, the intervals that appeared equal to 70°–75° became progressively larger as the anchor slant was decreased, and then leveled off as the adjusted intervals started to include both positive and negative slants. If the observers had been unbiased, or if

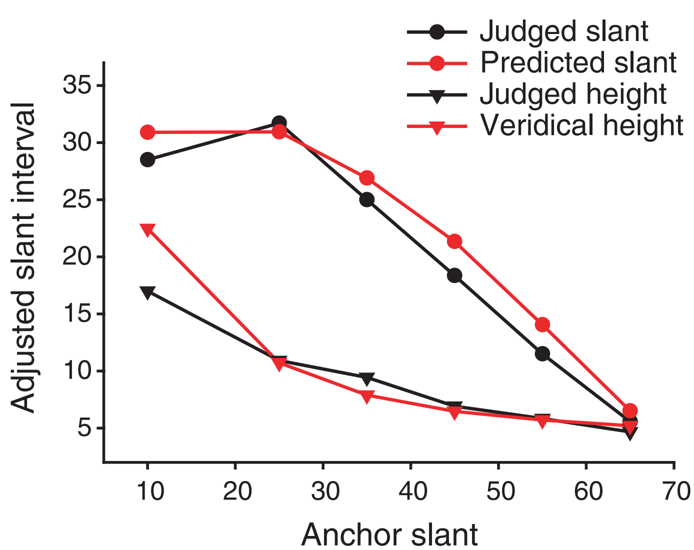

Figure 13. The average settings over all observers in Experiment 4. The black curves show the adjusted intervals as a function of the anchor slant for each task that appeared equal to the interval between 70° and 75°. The red curves show the predicted performance on the slant task based on scaling contrast and veridical performance on the tile height task.

their biases had been linear, then all of the adjusted intervals should have clustered around 5°. Clearly that did not happen. What the results show instead is that an interval anchored at 0° appears roughly six times smaller than an equivalent interval anchored at 70°.

The predicted pattern of performance based on the scaling contrast model is that the judged intervals should all have the same difference in scaling contrast as the 70°–75° standard that was used at the start of each experimental session. The adjusted intervals in the first block were all significantly smaller than what would be expected based on the model, which suggests that the observers may not have been able to resolve the full range of scaling differences in the 75° displays (see Todd et al., 2007). In order to correct for this, we estimated the effective scaling contrast between 70°–75° by computing the average difference in the adjustable interval of the first block over all observers and sessions. This estimate was then used to predict the judged intervals for all of the remaining 96 slant adjustment blocks in the entire experiment. The results of this analysis are shown in Figure 13 by the red curve with circular symbols, which closely tracks the mean adjusted intervals of the six observers.

A quite different pattern of results was obtained for the tile height adjustment task. Note in Figure 10 that the judged intervals for this task increased more slowly with decreasing anchor heights than is evident in the slant adjustment data. Moreover, these adjustments closely tracked the expected pattern of performance by an ideal observer who is capable of matching tile height intervals with perfect accuracy. This ideal response profile is shown in Figure 13 by the red curve with triangular symbols.

The problem of 2D cues

There is another interesting aspect of these data that deserves to be highlighted. If just noticeable differences of tile height and slant have similar profiles as the ones shown in Figure 13, then it is quite possible that observers may be more sensitive to changes in tile height than to changes in apparent slant. To better appreciate the potential significance of this suggestion, it is useful to consider a hypothetical experiment in which subjects are required to discriminate images of slanted surfaces whose presentations are paired with auditory tones whose intensities are systematically covaried with the depicted slants. If these tonal variations were sufficiently large, then subjects would almost certainly rely on them to maximize their discrimination performance. However, the results of such a study would tell us little or nothing about the basic perceptual mechanisms for estimating slants.

Of course no experienced researcher would ever knowingly employ such a design to investigate the perception of 3D structure, but there is a strong possibility that researchers who have used discrimination procedures to study the perception of slant from texture have done so inadvertently. Changes in slant inevitably produce changes in the projected heights and aspect ratios of optical texture elements. We know from the orthographic conditions of Experiment 1 and the earlier studies of Gillam (1970) and Todd et al. (2005, 2007) that projected heights and aspect ratios have little or no influence on the perceived slants of planar surfaces, yet the results of Experiment 4 suggest that observers may be more sensitive to changes in tile height than they are to changes in apparent slant. If that is indeed the case, then it is likely that they would rely primarily on height or shape changes to maximize performance in discrimination tasks, and the results would be no more informative about slant perception than the hypothetical experiment described above with covarying auditory tones.

Unfortunately, it is not possible to completely control these confounding 2D cues without introducing new ones. It is possible, however, to do a converse control: If observers are asked to discriminate orthographic images of slanted surfaces, the normal variations of projected heights and aspect ratios will be maintained, but the variations of apparent slant will be eliminated. If observers' discrimination thresholds for orthographic displays are equivalent to those obtained with perspective projections, it would provide compelling evidence that observers' judgments may not be based on perceived differences in slant. Experiment 5 was designed to examine this issue.

Experiment 5

Methods

Subjects

Five observers participated in the experiment, including two of the authors (JC and JT) and three others (DL, LT, and DW) who were naive about the issues being investigated. They all had normal or corrected-to-normal visual acuity.

Stimuli

There were three main conditions in which the displays depicted planar surfaces with different combinations of texture and perspective (see Figure 14). These included: (1) rounded tile textures with a 15° camera angle, (2) rounded tile textures under orthographic projection, and (3) horizontal stripe textures under orthographic projection. For each of these conditions, we created 1610 images that all had different randomizations of texture. The depicted planar surfaces had possible slants ranging from 0° to 80° in 0.5° increments, and 10 separate versions were created for each slant. Negative slants could be

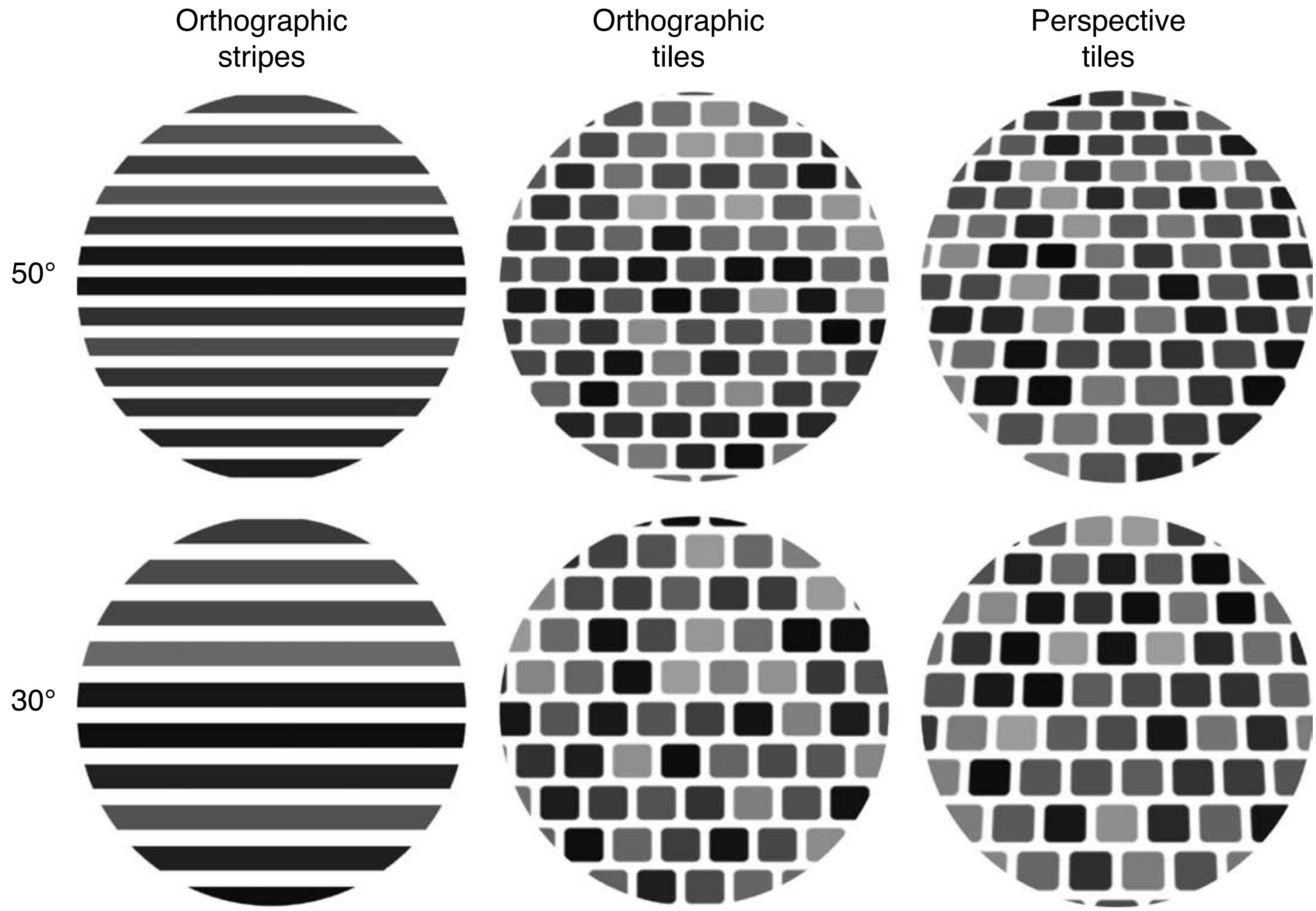

Figure 14. Some sample stimulus pairs from the different conditions of Experiment 5. All of the differences between the top and bottom rows are well above threshold. Readers should assess whether the differences in apparent slant in the perspective tile images are as salient as the differences in the heights or shapes of their respective texture elements (see also Figure 3).

created when needed by displaying an image with a positive slant upside down.

Procedure

Discrimination thresholds were measured using an "odd man out" paradigm, similar to the one employed by Hillis et al. (2002) in a cue combination study with texture and binocular disparity. There were four possible standard slants of 0°, 30°, 50°, or 70°. Each trial consisted of a sequence of three 1.5-s presentations of different images. These included two different versions of the standard with different randomizations of texture and a comparison stimulus that was presented once. Observers were required to indicate which stimulus was different from the other two on any basis by pressing an appropriate key on the computer keyboard. No feedback was given. Thresholds were estimated using a QUEST adaptive staircase procedure programmed in Matlab using the Psychophysics Toolbox extensions (Brainard, 1997). An experimental session included increment and decrement staircases for each of the different standards that were all interleaved with one another. Observers performed one session for each of the three combinations of texture and perspective.

Results and discussion

The average thresholds over all observers and staircase directions are shown in Figure 15 for each of the four standards with each combination of texture and perspective. It is important to keep in mind while considering these results that there were several possible sources of information that could have been used in order to identify the comparison stimulus on each trial. For the displays with tile textures under perspective projection, the observers' judgments could potentially have been based on differences in apparent slant or differences in the heights or aspect ratios of the optical texture elements. For the displays with tile textures under orthographic projection, there were no differences in apparent slant, but the standard and comparison stimuli could still be distinguished

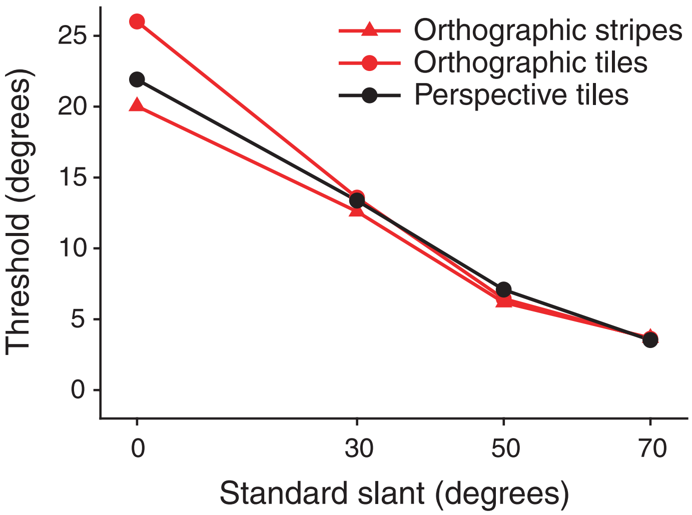

Figure 15. The average thresholds over all observers and staircase directions for the different combinations of texture and perspective used in Experiment 5.

by differences in projected height or aspect ratio. Finally, for the displays with striped textures under orthographic projection, the only relevant information for discriminating the stimuli was the projected heights of the stripes. What is especially striking in these data is that the thresholds were virtually identical in all three of these conditions. In other words, the observers' judgments were completely unaffected by whether or not the displays produced perceptible variations in apparent slant.

After each experimental session, we asked observers to provide a detailed description of the specific information they used to identify the comparison stimuli. They all confirmed that there were no variations in apparent slant for the displays rendered under orthographic projection, and that their responses to these stimuli were based on differences in the projected heights or shapes of the texture elements. We were particularly interested in their subjective impressions of the perspective tile condition. They all reported that they perceived differences in slant on some trials, especially those where the apparent slants were relatively large. However, they also indicated that there were other trials, in which they could not detect differences in apparent slant but could still identify the comparison stimuli based on differences in the projected heights or shapes of the texture elements. They reported that these trials occurred most frequently for the displays whose apparent slants were relatively small. Note that this is precisely the pattern that would be expected based on the subjective interval data shown in Figure 13. When considered in conjunction with the results of Experiment 4, these findings provide strong evidence that the discrimination of slanted planes from texture may often be based on 2D cues that are irrelevant to the perception of slant. If so, then the use of discrimination procedures may tell us little or nothing about how 3D slants are perceptually determined from patterns of optical texture.

General discussion

There have been many experiments reported in the literature that have employed discrimination procedures to estimate the variance of observers' slant judgments from texture and binocular disparity, both individually and in combination. One serious problem with this methodology, however, is that discrimination thresholds are influenced by several other factors in addition to variance, such as the slope of the psychometric function that relates physical and perceived slants. Thus, in order draw any conclusions about the variance of observers' slant estimates from discrimination data, it is necessary to assume that their slant estimates are unbiased—and to ignore the large literature that has documented biases in observers' perceptions over a wide variety of conditions (see Todd & Norman, 2003, for a review).

The results of the present experiments have demonstrated that the judged slants of planar surfaces from texture are systematically underestimated when observed with relatively small fields of view, as have typically been employed in previous discrimination studies. These results also demonstrate that perceived slant from texture varies as a curvilinear function of the depicted physical slant, in a manner that is consistent with the scaling contrast model originally proposed by Todd et al. (2007). The curvature of this function was revealed most clearly in Experiment 4, which showed that a slant interval anchored at 0°, where the slope is at its minimum, appears roughly six times smaller than an equivalent interval anchored at 70°, where the slope is much steeper.

Another serious problem with discrimination procedures for measuring the variance of observers' slant judgments is that the stimuli in these experiments inevitably contain 2D cues that can be used for successful discrimination performance but are unrelated to the perception of slant. For example, two such cues we have examined in detail are the heights and shapes of the optical texture elements. The results of Experiment 4 suggest that observers may actually be more sensitive to changes in these 2D cues than they are to changes in apparent slant. This was confirmed in Experiment 5 by comparing discrimination performance for displays generated using orthographic and perspective projections. The results revealed that observers' judgments were completely unaffected by whether or not the displays actually appeared slanted. Moreover, the observers in that study all reported that there were many trials in the perspective condition where they could not detect differences in apparent slant but could discriminate the stimuli nonetheless based on differences in the projected heights or shapes of the texture elements. It is obviously not possible to measure the variance of observers' slant estimates from judgments of some other property that is unrelated to the perception of slant.

Knill and Saunders (2003) recognized the problem of 2D cues in their experiments and attempted to mitigate it by adding random variations to the sizes of the texture elements in their displays by up to ±7%. Although this adds noise to the height cue, it does not eliminate its efficacy on average. It also has no effect on the 2D shape cue. This type of manipulation could be particularly informative if observers' performance on individual trials were correlated with the relative sizes of the texture elements for the standard and comparison stimuli, but we are unaware of any researchers who have performed such an analysis.

The inappropriate use of discrimination procedures has led to several questionable conclusions that have gained broad acceptance in the field. For example, Knill (1998a, 1998b) performed an extensive series of experiments and simulations from which he concluded that foreshortening cues (i.e., the aspect ratios of the optical texture elements) are the primary source of information for observers' judgments of slant. Although this was undoubtedly the primary source of information his observers used to discriminate the experimental displays, it is unlikely that these discriminations were based on the appearance of slant. The results obtained using orthographic projections in the present experiments show clearly that changes in foreshortening per se do not produce changes in the apparent slants of planar surfaces. The magnitude of apparent slant can be predicted quite accurately, however, by the magnitude of scaling contrast in an image. It is especially useful to highlight in this context an often overlooked article by Gillam (1970), who was the first to report that planar surfaces do not appear slanted in depth when foreshortening is the only available cue, and who also offered a prescient warning of how objective response data can easily be misinterpreted if one does not perform a careful assessment of observers' phenomenological impressions.

Another questionable conclusion from discrimination procedures that has been widely accepted in the literature is that the variance of observers' slant estimates from texture is an order of magnitude larger for shallow slants than for steep slants (e.g., see Hillis et al., 2004; Knill, 1998a, 1998b; Knill & Saunders, 2003). The variations in threshold as a function of slant obtained in these studies are almost certainly due to the curvilinear bias in observers' judgments, or to the use of 2D cues. When slant estimates were measured using adjustment tasks in the present experiments and the earlier studies of Todd et al. (2005, 2007), the standard deviations of observers' judgments remained approximately the same (5°–7°) over the entire range of slants investigated. The one exception to this was the set of images that appeared to have a frontoparallel orientation. The observers all commented that these were the judgments for which they had the greatest degree of confidence, and their standard deviations in those conditions were less than 2°.

A third questionable conclusion from discrimination procedures is that observers' judgments of surfaces with steep slants are more heavily influenced by texture than binocular disparity when the two cues are presented in combination (e.g., see Hillis et al., 2004; Knill & Saunders, 2003). We suspect that the relative cue weights in these studies may have been affected by factors other than estimated slant. It is important to keep in mind that the perceptual appearances of cue-combination stimuli have multiple attributes. Planar surfaces defined by stereo and texture can be perceived as slanted, but they also appear to be covered with a specific pattern of texture. For example, the variations in texture in Experiment 3 had a negligible effect on observers slant judgments, but they were clearly noticeable nonetheless and were perceptually interpreted as distortions of the surface texture. Although such distortions may not have been relevant to the appearance of slant, they certainly could be used as a source of information for discriminating stereograms.

If texture really is weighted more heavily than stereo in judgments of steep slants, then this ought to be evident from the data obtained using other experimental procedures that do not involve the measurement of discrimination thresholds. In order to assess this issue, it is useful to consider the results of other relevant studies that have examined the integration of texture and stereo using adjustment or magnitude estimation tasks, including Experiment 3 of the present series, and the earlier investigations of Buckley and Frisby (1993), Johnston et al. (1993), and Tittle et al. (1998). Although all of these studies included surfaces with steep slants, they all reached the same conclusion that the cue weights for texture are quite small relative to those for stereo. This pattern of results is theoretically quite sensible, given that judgments of depth or slant from stereo typically have much smaller systematic errors than comparable judgments of shape from texture. Because the differential bias for these cues is much larger than the standard deviation of observers' judgments, it follows that the optimal strategy to minimize error when both cues are combined is to rely primarily on stereo.

It is important to note in conclusion that the criticisms of prior studies we have outlined in this article are in no way directed at optimal cue integration theory as a potential model of human perception. Our criticisms are focused instead on one particular procedure that has been used in an effort to test that model. The hypothesis that observers combine information from multiple cues in a manner that minimizes perceptual errors is clearly attractive. We believe it is the case, however, that some researchers have construed the concept of error too narrowly by only considering the variance of observers' judgments, and not the much larger systematic biases that have been reported in numerous experiments on the perception of 3D structure.

Acknowledgments

This research was supported by a grant from the National Science Foundation (BCS-0546107). We are most grateful to Bart Anderson, Fulvio Domini, Barbara Gillam, and Jeff Saunders for their comments on earlier drafts of this manuscript.

Commercial relationships: none.
Corresponding author: James T. Todd.
Email: Todd.44@osu.edu.
Address: Department of Psychology, The Ohio State University, Columbus, OH 43210, USA.

Footnotes

[1]Because texture scaling is inversely related to the density of optical texture perpendicular to the direction of slant, scaling contrast is mathematically equivalent to density contrast. Thus, another reasonable interpretation of Todd et al.'s (2007) data is that observers' judgments may have been based on systematic variations in texture density or spatial frequency among different local regions of an image (see also Thaler, Todd, & Dijkstra, 2007; Todd et al., 2005).

[2]In order to confirm the validity of this procedure, we compared stereograms composed of orthographic images of surfaces with distorted back-projected textures to those that were rendered under perspective projection with homogeneous tile textures. If the back-projected textures are scaled appropriately, and depict the same slant as specified by binocular disparity, then the stereograms generated with these two procedures are completely indistinguishable.

References

Aloimonos, J. (1988). Shape from texture. *Biological Cybernetics, 58,* 345–360. [PubMed]

Backus, B. T., & Saunders, J. A. (2006). The accuracy and reliability of perceived depth from linear perspective as a function of image size. *Journal of Vision, 6*(9):7, 933–954, http://journalofvision.org/6/9/7/, doi:10.1167/6.9.7. [PubMed] [Article]

Bajcsy, R., & Lieberman, L. (1976). Texture gradient as a depth cue. *Computer Graphics Image Processing, 5,* 52–67.

Blake, A., & Marinos, C. (1990). Shape from texture: Estimation, isotropy and moments. *Artificial Intelligence, 45,* 323–380.

Blostein, D., & Ahuja, N. (1989). Shape from texture: Integrating texture-element extraction and surface estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 11,* 1233–1251.

Bradshaw, M. F., Parton, A. D., & Glennerster, A. (2000). The task-dependent use of binocular disparity and motion parallax information. *Vision Research, 40,* 3725–3734. [PubMed]

Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision, 10,* 433–436. [PubMed]

Braunstein, M. L., & Payne, J. W. (1969). Perspective and form ratio as determinants of relative slant judgments. *Journal of Experimental Psychology, 81,* 584–590. [Article]

Brown, L. G., & Shvayster, H. (1990). Surface orientation from projective foreshortening of isotropic texture autocorrelation. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 12,* 584–588.

Buckley, D., & Frisby, J. P. (1993). Interaction of stereo, texture, and outline cues in the shape perception of three-dimensional ridges. *Vision Research, 33,* 919–934. [PubMed]

Clerc, M., & Mallat, S. (2002). The texture gradient equation for recovering shape from texture. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 24,* 536–549.

Domini, F., & Caudek, C. (2010). Matching perceived depth from disparity and from velocity: Modeling and psychophysics. *Acta Psychologica, 133,* 81–89.

Ernst, M. O., & Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature, 415,* 429–433. [PubMed]

Gårding, J. (1992). Shape from texture for smooth curved surfaces in perspective projection. *Journal of Mathematical Imaging and Vision, 2,* 327–350. [Article]

Gårding, J. (1993). Shape from texture and contour by weak isotropy. *Artificial Intelligence, 64,* 243–297. [Article]

Gillam, B. (1968). Perception of slant when perspective and stereosis conflict: Experiments with aniseikonic lenses. *Journal of Experimental Psychology, 78,* 299–305. [PubMed]

Gillam, B. (1970). Judgments of slant on the basis of foreshortening. *Scandinavian Journal of Psychology, 11,* 31–34. [PubMed]

Hawkins, D. M., & Wixley, R. A. J. (1986). A note on the transformation of chi-squared variables to normality. *The American Statistician, 40,* 296–298.

Hillis, J. M., Ernst, M. O., Banks, M. S., & Landy, M. S. (2002). Combining sensory information: Mandatory fusion within, but not between, senses. *Science, 298,* 1627–1630. [PubMed]

Hillis, J. M., Watt, S. J., Landy, M. S., & Banks, M. S. (2004). Slant from texture and disparity cues: Optimal cue combination. *Journal of Vision, 4*(12):1, 967–992, http://journalofvision.org/4/12/1/, doi:10.1167/4.12.1. [PubMed] [Article]

Johnston, E. B., Cumming, B. J., & Parker, A. J. (1993). Integration of depth modules: Stereopsis and texture. *Vision Research, 33,* 813–826. [PubMed]

Knill, D. C. (1998a). Discriminating surface slant from texture: Comparing human and ideal observers. *Vision Research, 38,* 1683–1711. [PubMed]

Knill, D. C. (1998b). Ideal observer perturbation analysis reveals human strategies for inferring surface orientation from texture. *Vision Research, 38,* 2635–2656. [PubMed]

Knill, D. C., & Saunders, J. (2003). Do humans optimally integrate stereo and texture information for judgments of surface slant? *Vision Research, 43,* 2539–58. [PubMed] [Article]

Koenderink, J. J., & Van Doorn, A. J. (1996). Illuminance texture due to surface mesostructure. *Journal of the Optical Society of America A, 13,* 452–463.

Krumm, J., & Shafer, S. A. (1992). Shape from periodic texture using the spectrogram. *Proceedings of the 1992 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR '92),* 284–289. [Article]

Malik, J., & Rosenholtz, R. (1994). A computational model for shape from texture. In G. R. Bock & J. A. Goode (Eds.), *Higher order processing in the visual system* (pp. 272–286). Chichester, UK: John Wiley and Sons.

Malik, J., & Rosenholtz, R. (1997). Computing local surface orientation and shape from texture for curved surfaces. *International Journal of Computer Vision, 23,* 149–168. [Article]

Marinos, C., & Blake, A. (1990). Shape from texture: The homogeneity hypothesis. *Proceedings of the 3rd International Conference in Computer Vision, Osaka, Japan,* 350–334. [Article]

Purdy, W. P. (1958). The hypothesis of psychophysical correspondence in space perception. *Dissertation Abstracts International, 42,* 1454. (University Microfilms No. 58–5594).

Sakai, K., & Finkel, L. H. (1995). Characterization of spatial frequency spectrum in the perception of shape from texture. *Journal of the Optical Society of America A, Optics, Image Science, and Vision, 12,* 1208–1224. [PubMed]

Super, B. J., & Bovik, A. C. (1995). Planar surface orientation from texture spatial frequencies. *Pattern Recognition, 28,* 729–743. [Article]

Thaler, L., Todd, J. T., & Dijkstra, T. M. H. (2007). The effects of phase on the perception of 3D shape from texture: Psychophysics and modeling. *Vision Research, 47,* 411–427. [PubMed]

Tittle, J. S., & Braunstein, M. L. (1993). Recovery of 3-D shape from binocular disparity and structure from motion. *Perception & Psychophysics, 54,* 157–169. [PubMed]

Tittle, J. S., Norman, J. F., Perotti, V. J., & Phillips, F. (1998). The perception of scale-dependent and scale-independent surface structure from binocular disparity, texture, and shading. *Perception, 27,* 147–166. [PubMed]

Tittle, J. S., & Perotti, V. J. (1997). The perception of shape and curvedness from binocular stereopsis and structure from motion. *Perception & Psychophysics, 59,* 1167–1179. [PubMed]

Tittle, J. S., Todd, J. T., Perotti, V. J., & Norman, J. F. (1995). The systematic distortion of perceived 3D structure from motion and binocular stereopsis. *Journal of Experimental Psychology: Human Perception and Performance, 21,* 663–678. [PubMed]

Todd, J. T., & Norman, J. F. (2003). The perception of 3D shape from multiple cues: Are observers capable of perceiving metric structure? *Perception & Psychophysics, 65,* 31–47. [PubMed]

Todd, J. T., Oomes, A. H. J., Koenderink, J. J., & Kappers, A. M. L. (2004). The perception of doubly curved surfaces from anisotropic textures. *Psychological Science, 15,* 40–46. [PubMed]

Todd, J. T., & Perotti, V. J. (1999). The visual perception of surface orientation from optical motion. *Perception & Psychophysics, 61,* 1577–1589. [PubMed]

Todd, J. T., Thaler, L., & Dijkstra, T. M. H. (2005). The effects of field of view on the perception of 3D slant from texture. *Vision Research, 45,* 1501–1517. [PubMed]

Todd, J. T., Thaler, L., Dijkstra, T. M. H., Koenderink, J. J., & Kappers, A. M. L. (2007). The effects of viewing angle, camera angle and sign of surface curvature on the perception of 3D shape from texture. *Journal of Vision, 7*(12):9, 1–16, http://journalofvision.org/7/12/9/, doi:10.1167/7.12.9. [PubMed] [Article]

Van Ee, R., & Erkelens, C. J. (1996). Temporal aspects of binocular slant perception. *Vision Research, 36,* 43–51. [PubMed]

Watt, S. J., Akeley, K., Ernst, M. O., & Banks, M. S. (2005). Focus cues affect perceived depth. *Journal of Vision, 5*(10):7, 834–862, http://journalofvision.org/5/10/7/, doi:10.1167/5.10.7. [PubMed] [Article]

Witkin, A. P. (1981). Recovering surface shape and orientation from texture. *Artificial Intelligence, 17,* 17–45.