Perceptual biases and cue weighting in perception of 3D slant from texture and stereo information

Jeffrey A. Saunders

Zhongting Chen

Multiple cues are typically available for perceiving the 3D slant of surfaces, and slant perception has been used as a test case for investigating cue integration. Previous evidence suggests that texture and stereo slant cues contribute in an optimal Bayesian manner. We tested whether a Bayesian model could also account for perceptual underestimation of slant from texture. One explanation proposed by Todd, Christensen, and Guckes (2010) is that slant from texture is based on an inaccurate optical variable. An alternative Bayesian explanation is that perceptual underestimation is due to the influence of frontal cues and/or a frontal prior, which is weighted according to the reliability of slant cues. We measured slant perception using a hand-alignment task for conditions that provided only texture, only stereo, or combined texture and stereo cues. Slant estimates from monocular texture showed large biases toward frontal, with proportionally more underestimation at low slants than high slants. Slant estimates from stereo alone were more accurate, and adding texture information did not reduce accuracy. These results are consistent with a frontal influence that is decreasingly weighted as slant information becomes more reliable. We also included conditions with small cue conflicts to measure the relative weighting of texture and stereo cues. Consistent with previous studies, texture had a significant effect on slant estimates in binocular conditions, and the relative weighting of texture increased with slant. In some cases, perceived slant from combined stereo and texture cues was higher than from either cue in isolation. Both the perceptual biases and the cue weights were generally consistent with a Bayesian model that optimally integrates texture and stereo slant cues with frontal cues and/or a frontal prior.

Introduction

There are many sources of 3D information available, and the problem of how 3D cues are integrated is an active area of research. Evidence from a number of recent studies suggests that the visual system makes efficient use of available information and combines cues in an optimal manner (e.g., Jacobs, 1999; Ernst & Banks, 2002; Knill & Saunders, 2003; Hillis, Watt, Landy, & Banks, 2004; Saunders & Backus, 2006a).

A number of cue integration studies have investigated the specific case of stereo and texture information about 3D slant (Buckley & Frisby, 1993; Knill & Saunders, 2003; Hillis et al., 2004; Watt, Akeley, Ernst, & Banks, 2005; Girshick & Banks, 2009). One reason that perception of slant from texture is an interesting test case is because the reliability of texture information changes dramatically as a function of slant (Knill, 1998a). Consistent with this difference in information, observers show high discrimination thresholds when comparing surfaces that are near frontal, and much lower discrimination thresholds when comparing surfaces are highly slanted relative to the observer (Knill, 1998b; Knill & Saunders, 2003; Hillis et al., 2004; Rosas, Wichmann, & Wagemans, 2004). Optimal integration therefore predicts that the relative weighting of texture information would strongly vary as a function of slant. The empirical results observed by Knill and Saunders (2003) and Hillis et al. (2004) were consistent with this prediction. Both studies found that weighting of texture increased with slant, and further that the change in cue weights was quantitatively consistent with measured reliabilities of stereo and texture slant cues. These results suggest that stereo and texture slant information are integrated in an optimal manner.
Perceptual underestimation of slant from texture

A number of previous studies have found that perception of slant from monocular texture is biased toward the frontal plane, with proportionally more underestimation at low slants than high slants (Todd, Thaler, & Dijkstra, 2005; Watt et al., 2005; Norman et al., 2009; Durgin, Li, & Hajnal, 2010; Todd et al., 2010). Figure 1 shows examples of perspective views of a textured surface slanted by 30° and 50° relative to the frontal plane. The Voronoi texture is isotropic and homogeneous, so the projected texture gradient could in principle allow slant to be accurately recovered. However, in the case of the surface with 30° slant (left), the surface appears almost frontal. The surface with 50° slant produces a clearer percept of a surface slanted in depth, but slant is still underestimated. Based on previous results, the perceived slant would be about half of the veridical slant for a stimuli with this field of view (Todd et al., 2005; Todd et al., 2010).

An explanation proposed by Todd et al. (2005) and Todd, Thaler, Dijkstra, Koenderink, and Kappers (2007) is that perceived slant from texture is based on an inaccurate optical variable: the scaling contrast of texture elements across an image. This variable was identified based on empirical observations. Todd et al. (2005) compared human judgments of the slant of textured surfaces with various possible measures that could be computed from a projected image. They found that the overall change in the size and density of texture across a projected image was strongly correlated with slant judgments across conditions with varying slant and field-of-view, and that this variable was a better predictor than the actual simulated slant. Further results by Todd et al. (2007) led them to propose a more specific variable,

\[ \text{Scaling contrast} = \frac{\lambda_{\text{max}} - \lambda_{\text{min}}}{\lambda_{\text{max}} + \lambda_{\text{min}}}, \]  

where \( \lambda_{\text{max}} \) and \( \lambda_{\text{min}} \) are the maximum and minimum sizes of texture elements in the projected image as measured in the direction perpendicular to the depth gradient (i.e., the widths in Figure 1). For a planar surface with homogeneous texture, the scaling contrast can be expressed in terms of the surface slant and field of view:

\[ \text{Scaling contrast} = \tan(S)\tan(F/2), \]  

where \( S \) is slant relative to the line of sight at the center of an image and \( F \) is the diameter of the field-of-view. Todd et al. (2007) and Todd et al. (2010) observed an approximately linear relationship between scaling contrast and perceptual judgments of slant, in contrast to the nonlinear relationship between perceptual judgments and actual slant. They propose that use of scaling contrast as a slant cue could explain the perceptual biases observed for judgments of slant from texture.

Perceptual underestimation of slant could also arise due to the influence of a prior or conflicting cues specifying a frontal orientation, even if visual processing of texture information were unbiased. The virtual surfaces used in most psychophysical studies provide
conflicting accommodative blur cues. When a slanted virtual surface is presented with a frontal display screen, the accommodative cues specify the frontal orientation of the display surface rather than the simulated orientation of the virtual surface. The presence or absence of a blur gradient has been shown to affect slant perception (Watt et al., 2005; Norman et al., 2009). If texture information is relatively weak, it might not be sufficient to fully counteract the influence of conflicting information from accommodative blur cues. Another factor could be a general bias toward uniform depth (Gogel, 1965; Hillis et al., 2004). A general implicit bias has been posited to explain underestimation of depth in situations when visual information is limited (e.g., Bülthoff & Mallot, 1988; Ooi, Wu, & He, 2006; Wu, Ooi, & He, 2004; Saunders & Backus, 2006b). In a Bayesian framework, this could be modeled as a prior distribution that contributes like an additional sensory cue. For slant perception, the effect of such a prior would be similar to that of conflicting sensory cues specifying a frontal surface, leading to underestimation of slant when other slant cues are weak.

Thus, there are at least two qualitatively different explanations for biased perception of slant from texture. Texture information might be interpreted in a biased manner, as suggested by Todd et al. (2007) and Todd et al. (2010), or might be accurately interpreted but too weak to counteract the influence of other cues or prior assumptions. These possibilities can be distinguished by how the texture affects perceived slant when other sources of information like binocular disparity are available, as described in the next section.

**Bayesian model of frontal bias**

In a Bayesian model, different sources of information are represented as likelihood functions that are combined with a prior distribution to estimate a posterior probability. The relative influence of different cues on the posterior function is dependent on the spread of the likelihood functions from each cue. Figure 2 illustrates the combination of conflicting information from texture and stereo cues in a situation where texture information is less reliable than stereo information. The likelihood functions from texture and stereo have different peaks corresponding to the slants specified by each cue. The spread of the likelihood functions represents the range of slants that would be consistent with the cues. In this example, texture information is less reliable, so the likelihood function has a wider distribution. When the two likelihood functions are multiplied and normalized, the resulting combined likelihood function is narrower and has an intermediate peak. Due to the difference in reliability of cues, the maximum of the combined likelihood function is closer to the slant specified by stereo than the slant specified by texture. If one adopts an a priori assumption that all possible slants are equally likely (i.e., a uniform prior), the slant with maximum likelihood corresponds to the slant with highest posterior probability given the available information. In the case where both likelihood functions are Gaussian, the maximum likelihood estimate of slant can be expressed as a weighted average of the slants specified by each cue, $\hat{S}_{\text{st+tex}} =$
$w_{\text{tex}} S_{\text{tex}} + w_{\text{ster}} S_{\text{ster}}$, with weights proportional to the reliability of the cues, $w_{\text{tex}}/w_{\text{ster}} = (1/\sigma_{\text{tex}}^2)/(1/\sigma_{\text{ster}}^2)$ (see Ernst & Banks, 2002 or Knill & Saunders, 2003).

We now consider the effect of conflicting cues specifying a frontal surface, such as accommodative cues when artificial stimuli are presented on a frontal display. The information from conflicting frontal cues can be modeled as an additional likelihood function with peak at zero, which is multiplied with the likelihood functions from texture and/or stereo cues. The effect would be to bias perceived slant toward frontal, and the amount of bias would depend on the reliability of the other cues. Figure 2b illustrates the effect of combining frontal cues with an unreliable texture cue specifying nonzero slant. In this example, the estimated likelihood function from texture has a peak at the veridical slant, but has a wide distribution reflecting the low reliability of texture information. Even though the estimated likelihood function from texture is unbiased, the maximum of the combined distribution shows a large shift toward zero due to the influence of the conflicting frontal cues. Figure 2c illustrates the effect of combining the same frontal cues with more reliable information from stereo. The maximum of the posterior is again shifted toward zero, but the amount of bias is smaller because the likelihood distribution from stereo information is more tightly constrained. Figure 2d illustrates the case of combining frontal cues with both texture and stereo information. The combined texture and stereo cues are more reliable than either cue individually, so the conflicting frontal cues have less influence. Note that slant estimate from combined texture and stereo (Figure 2d) is higher than the slant estimates from either texture or stereo alone (Figures 2b, 2c). This is due to the influence of the frontal cues. Accurate estimation of slant requires not only accurate interpretation of texture and stereo cues, but also that these cues are sufficiently reliable to overcome the influence of the conflicting frontal cues.

Combining visual cues with a nonuniform prior could have a similar effect as combining with conflicting frontal cues. For slant estimation, a Bayesian prior would represent the general probability of different slants in the absence of any visual cues about the specific situation. If some slants tend to be experienced more frequently than others, such as frontal surfaces, then the appropriate prior distribution would not be uniform. For example, Hillis et al. (2004) suggest a prior for slant that is proportional to cos(S), which would be the distribution of local slant in a projected image if surface orientations were uniformly distributed in 3D space. If the visual system incorporates prior knowledge that low slants are more likely than high slants when estimating slant from visual cues, the result would be underestimation of slant. In a Bayesian model, priors are multiplied with likelihood functions to determine the posterior probability, and therefore contribute like an additional visual cue. In the examples shown in Figures 3b through d, the dashed lines can be interpreted as representing a prior distribution rather than the likelihood function from conflicting visual cues, and the results would be equivalent. In this study, we will not attempt to distinguish between the effect of conflicting frontal cues and a nonuniform prior, but rather treat these as two potential sources of a general frontal bias. In a Bayesian model, the influence of either factor would tend to bias perceived slant toward frontal by an amount dependent on the reliability of visual slant cues.

This model could account for the nonlinear pattern of perceptual underestimation of slant from texture observed in previous studies. Texture is a more informative slant cue when slant is high (Knill, 1998a), so frontal cues or a frontal prior would have less influence at high slants than at low slants. The proportional amount of perceptual underestimation would therefore vary as a function of slant, resulting in a nonlinear psychometric function. As a measure of perceptual underestimation, we will define perceptual gain to be the proportional amount of change in perceived slant caused by a change in the slant specified by texture and/or stereo cues. The perceptual gain can be thought of as the weighting of slant cues relative to a frontal tendency. At low slants, $\sigma_{\text{tex}}$ would be relatively large, resulting in a low perceptual gain. At higher slants, $\sigma_{\text{tex}}$ gets much smaller and the perceptual gain would approach unity. A nonlinear psychometric function would arise because of the differences in the reliability of texture information at different slants. An accurate model of the likelihood function from texture would be asymmetric and skewed toward zero (Knill, 1998a), and the likelihood function from frontal cues might be non-Gaussian as well. However, the qualitative prediction would be similar regardless of the exact form of the distributions: systematic increase in perceptual gain as the slant specified by texture is increased.

By comparing slant estimates in single cue and combined cue conditions, this model of perceptual underestimation of slant can be distinguished from a model in which texture information is interpreted in a biased manner. If underestimation were due to the influence of a general frontal tendency, then increasing the reliability of slant information would tend to make perceptual estimates more accurate. Even if perceptual estimates of slant from texture alone were highly biased, adding texture to stereo information would result in more accurate slant estimates than from stereo information alone. It would be possible for perceived slant from combined stereo and texture cues to be higher than perceived slant from either stereo or texture cues in isolation. In contrast, if perceptual underesti-
mation of slant from texture alone were due to biased interpretation of texture information, then any influence of texture in combined cue conditions would reduce accuracy. If texture contributes to perceived slant when stereo is available, which can be assessed from cue conflict conditions, then the effect would be to reduce perceived slant compared to conditions with stereo only. These contrasting predictions are tested in the experiments reported here.

More generally, the Bayesian model predicts that cue reliability would determine both the relative weighting of slant cues and the amount of perceptual underestimation. A frontal tendency would contribute as an additional cue that is constant across conditions, so the relative weighting of this “cue” would depend on the reliability of texture and/or stereo information. At low slants, when texture is less reliable than stereo, one would expect low weighting of texture relative to stereo and lower perceptual gain from texture only than from stereo only. At higher slants, when texture information is more reliable, one would expect both greater weighting of texture relative to stereo and higher perceptual gain in texture only conditions.

With some additional assumptions, the relative reliability of stereo and texture information could be inferred from perceptual biases in the single cue conditions, which could then be used to predict optimal cue weights. Suppose that the combined influence of frontal cues and/or a frontal prior can be represented by a Gaussian distribution centered at zero with width $\sigma_o$, and that the likelihood functions from stereo and texture are also Gaussian distributions with widths $\sigma_{\text{ster}}$ and $\sigma_{\text{tex}}$. Assuming optimal Bayesian integration, the perceptual gain in single cue conditions would be:

$$g_{\text{tex}} = \sigma_0^2/(\sigma_{\text{tex}}^2 + \sigma_0^2), \quad g_{\text{ster}} = \sigma_0^2/(\sigma_{\text{ster}}^2 + \sigma_0^2).$$

These equations can be combined to solve for the ratio of reliabilities,

$$\frac{\sigma_{\text{tex}}^2}{\sigma_{\text{ster}}^2} = \frac{1}{g_{\text{tex}}} - 1 / \left(1 - \frac{1}{g_{\text{ster}}} - 1\right).$$

If the relative weighting of stereo and texture in combined cue conditions were optimal, then the ratio of reliabilities $\sigma_{\text{tex}}^2 / \sigma_{\text{ster}}^2$ would be equal to the ratio of cue weights $w_{\text{ster}} / w_{\text{tex}}$. We tested whether the relative reliabilities computed from perceptual biases using Equation 4 were consistent with measures of cue weights derived from cue conflict conditions.

**Estimation of texture and stereo cue weights**

To estimate the relative weighting of stereo and texture cues for slant perception, the previous studies by Knill and Saunders (2003) and Hillis et al. (2004) used a discrimination task and cue conflict stimuli, as illustrated in Figure 3. Observers discriminated the slant of surfaces with and without small amounts of conflicting slant information. For one of the two surfaces on a discrimination trial, stereo and texture information specified different slants. These cue conflict stimuli were compared to reference surfaces with consistent stereo and texture information. The slant of the reference surface was varied across trials to determine a point of subjective equality (PSE). The relative weighting of stereo and texture information can be inferred from whether the PSE was close to the slant specified by stereo or the slant specified by texture.

Todd and Christensen, and Guckes (2010) have raised a concern about the method used to estimate cue weights in these previous studies. They argue that observers might have been performing discrimination judgments based on 2D image cues rather than an integrated perception of 3D slant. Specifically, observers might compare the amount of foreshortening of texture in a projected image across pairs of stimuli. Use of this strategy could result in apparent weighting of texture information even if the texture cue was not actually integrated with stereo information to perceive 3D slant.

Results of Hillis, Ernst, Banks, and Landy (2002) suggest that observers cannot independently access texture and stereo cues when viewing binocular images of a slanted textured surface, which would alleviate the concern about use of 2D strategies in slant discrimination. Hillis et al. used an odd-man-out task that could potentially be performed based on texture and stereo information used independently, and therefore did not require an integrated perception of slant.

Performance in some key conditions was not as good as would be expected if observers were able to separately attend to texture and stereo cues. Hillis et al. (2002) interpret their results as evidence for mandatory fusion of texture and stereo cues. If this finding generalizes to the conditions tested by Knill and Saunders (2003) and Hillis et al. (2004), then the measured cue weights would be valid indicators of the relative influence of stereo and texture cues on perception of 3D slant.

On the other hand, Todd et al. (2010) tested slant estimation for cue conflict stimuli and found no apparent weighting of texture information. They presented binocular stimuli with various conflicts
between stereo and texture slant information, and observers used a gauge figure to indicated perceived 3D slant. Todd et al. found that performance was dominated by the slant specified by stereo. Even at high slants, for which texture information produces reliable discrimination judgments, Todd et al. did not observe a strong influence of texture. The lack of influence from texture is contrary to the findings of Knill and Saunders (2003) and Hillis et al. (2004). Todd et al. (2010) suggest that the previous results were an artifact of using a discrimination task, as described above.

However, a significant difference between these studies is that Todd et al. (2010) used much larger cue conflicts, 20° and above. When large conflicts are present, the visual system might employ a robust strategy of vetoing less reliable cues (Landy, Maloney, Johnston, & Young, 1995), or use a mixture model to interpret depth cues in a robust manner (Knill, 2003). A number of studies have observed effects consistent with robust integration of monocular and binocular slant cues (Banks & Backus, 1998; van Ee, van Dam, & Erkelens, 2002; Knill, 2007a; Girshick & Banks, 2009). In particular, Girshick and Banks (2009) tested perception of slant from stereo and texture cues for stimuli with varied amounts of cue conflicts. They found that stereo and texture cues with small discrepancies were integrated to perceive an intermediate slant, but with larger cue conflicts observers tended to rely on only one of the two cues. The large cue conflicts used in Todd et al. (2010) could have similarly elicited robust integration, which would explain why Todd et al. did not find any influence of texture information. The large cue conflicts could also have encouraged a reliance on stereo information. Some studies have demonstrated that cue weighting can be dynamically modified through training (Ernst, Banks, & Bulthoff, 2000; Knill, 2007b). In Todd et al.’s (2010) study, frequent exposure to stimuli with large cue conflicts could have

Figure 3. Illustration of how the relative weighting of stereo and texture cues can be measured using cue conflict stimuli. For the test surface (top), stereo information specifies a slant of 50° while the texture gradient specifies a slant of 55°. The perceived slant of this test surface can be compared to that of a reference surface with consistent stereo and texture information. By varying the slant of the reference surface across trials, one can find a point of subjective equality. Suppose that the test surface is perceived to have the same slant as a reference surface with stereo and texture slant of 52° (bottom). This implies that increasing the slant specified by texture by 5° has the same effect as increasing both stereo and texture slant by 2°, and conversely that decreasing the slant specified by stereo by 5° has the same effect as decreasing both stereo and texture slant by 3°. The corresponding cue weights would be 0.4 for texture and 0.6 for stereo in this example.
encouraged or trained observers to rely on stereo more than they would have otherwise.

In the present study, we used a slant estimation task to assess perceived slant in cue conflict conditions, but tested much smaller cue conflicts than Todd et al. (2010). Slant varied from $5^\circ$ to $60^\circ$ and conflicts were limited to $\pm 5^\circ$. This is more comparable to the size of conflicts tested in the previous studies that found optimal integration of texture and stereo slant cues rather than cue vetoing (Knill & Saunders, 2003; Hillis et al., 2004). The relative weighting of texture and stereo information was estimated in the same manner as illustrated in Figure 3, except that estimates of slant from an adjustment task were used rather than PSEs from discrimination judgments.

We used a hand alignment task to measure slant perception rather than a 2D gauge figure task like used by Todd et al. (2010). Our method was similar to that of Durgin et al. (2010): Observers wore a flat board on palm of their right hand and adjusted the orientation of their hand to be aligned with a presented surface. We also tested a 2D gauge figure task in a pilot experiment, but found that responses of naïve observers often showed a large amount of compression even in full cue conditions, and there were large individual differences in the scaling of responses. For the same conditions, the hand alignment task used here was found to produce more accurate slant judgments with more consistent scaling. However, a disadvantage of this task is that it depends on a mapping from visual coordinates to hand coordinates, which could introduce additional measurement errors. In particular, there were errors in the orientation of the hand when matched to a frontal surface, which are likely due to the task rather than perceptual bias. We normalized responses to remove constant bias when computing measures of perceptual gain, and estimated the relative weighting of texture and stereo cues in a manner that would be invariant to both constant and multiplicative response bias.

To estimate cue weights, we compared the effect of perturbing individual cues with the effect of changing both texture and stereo in a consistent manner. Slant judgments from consistent texture and stereo were obtained at a range of base slants. The slope of the psychometric function at each base slant, $b_{st+tex}$, indicates the change in hand orientation that would result from changing the slant specified by both texture and stereo cues by a consistent amount. At various base slants, we also tested cue conflict conditions in which the slant specified by texture was slightly larger or smaller than the slant specified by stereo. From these results, one can measure a slope representing the effect of changing just the texture cue, $b_{tex}$. The ratio $w_{tex} = b_{tex} / b_{st+tex}$ represents the effect of changing the texture cue as a proportion of the effect of changing both cues, which provides a measure of the relative weighting of texture information. Because $w_{tex}$ is a ratio of slopes, it would be invariant to any scaling due to response bias. In a situation where perceived slant was entirely determined by texture information, the slopes would be equal and the texture weight would be $w_{tex} = 1$. In a situation where perceived slant were entirely determined by stereo information, $b_{tex}$ would be zero and consequently $w_{tex} = 0$. As an intermediate example, suppose that the slopes were found to be $b_{st+tex} = 0.75$ and $b_{tex} = 0.30$. The texture weight implied by these slopes would be $w_{tex} = 0.30/0.75 = 0.4$. For comparison, the slopes can be used to predict the results of an experiment using a discrimination paradigm. Increasing the slant specified by texture by $5^\circ$ would be expected to have the same effect $0.3 \times 5^\circ = 1.5^\circ$ as increasing slant specified by both the stereo and texture slants by $2^\circ$ ($0.75 \times 2^\circ = 1.5^\circ$). The texture weight implied by these matching conditions would be $w_{tex} = 2^\circ/5^\circ = 0.4$, which is equal to the ratio of $b_{tex}$ to $b_{st+tex}$. Thus, our analysis allows cue weights to be computed from direct estimates of perceived slant in a manner that is analogous to the analyses of Knill and Saunders (2003) and Hillis et al. (2004).

**Figure 4.** Broadband noise texture used to provide disparity information in binocular conditions without providing an effective monocular slant cue. The images show surfaces slanted by $30^\circ$ and $50^\circ$ relative to the frontal plane. In contrast to the surfaces with Voronoi texture shown in Figure 1, neither of these images appears to be slanted in depth.

---

**Experiment 1**

Experiment 1 measured slant perception using a hand alignment task for stimuli with varied slant information. Two surface textures were tested: a Voronoi texture that produces a strong monocular cue to slant (Figure 1) and a broadband noise texture that does not (Figure 4). Slanted surfaces were presented both monocularly and binocularly, and with and without cue conflicts. Based on previous results, we
expected that the Voronoi texture would affect perceived slant in binocular conditions, and that the weighting of the texture cue would increase with slant. A further prediction of the Bayesian model is that slant underestimation would vary with the reliability of slant information. Because the slant information provided by texture improves with slant, one would expect proportionally less underestimation at higher slants for the Voronoi texture viewed under monocular viewing conditions. Binocular stimuli with the Voronoi texture, which provide both stereo and texture information about slant, would be expected to produce less underestimation than binocular stimuli with broadband texture, which provide only stereo information.

Methods
Participants
Twelve adults (seven males and five females) at University of Hong Kong participated in the experiment. One participant was excluded because he showed highly discrepant response patterns in the two experimental sessions. All the participants had normal or corrected-to-normal visual acuity, and passed a stereo acuity screening test. All participants were naive as to the purpose of the study and were paid for participating. The procedures were approved by and conform to the standards of the Human Research Ethics Committee for Non-Clinical Faculties.

Apparatus and stimuli
The stimuli were computer-generated perspective images of the slanted planar surfaces viewed through a 16° diameter circular aperture on a black mask. The images were presented on a LCD monitor (ASUS VG278H) that had a resolution of 1920 × 1080 pixels and a refresh rate of 120 Hz (60 Hz for each eye). The display had a 59.2 cm × 33.6 cm viewable region and was viewed from a chin rest at a distance of 100 cm. Shutter glasses (NVIDIA 3D Vision 2) were used to present left and right stereo images to the two eyes. Interocular distance was measured for each individual participant and used to compute accurate stereo images. For monocular viewing conditions, the non-dominant eye of a participant was covered with an eye patch. Images were rendered with OpenGL using a NVIDIA Quadro 600 graphics card, and were anti-aliased with subpixel resolution.

Observers indicated the perceived slant of a surface by aligning their hand. To record the position of the hand, we used a 3D Guidance trakSTAR system. Observers wore a board on their palm that had three sensors that tracked the 3D positions. The orientation of the palm was computed from the locations of the three sensors.

Two types of surface texture were used. The Voronoi texture (Figure 1) was similar to that used in many previous studies of slant-from-texture (e.g., Knill, 1998b; Knill & Saunders, 2003; Hillis et al., 2004; Saunders & Backus, 2006a), and is known to provide an effective slant-from-texture cue even under monocular viewing conditions. To determine the center points of a Voronoi pattern, we first generated uniformly distributed random points on a square region and then applied a repulsion process to make the spacing of center points more regular. We generated tile-able square patches of Voronoi texture that had 222 cells per tile, and then tiled these patches to extend the texture over the visible region of a surface. The scale of the texture was randomly varied from trial to trial so that the projected size of texture elements did not provide a reliable cue to surface slant. At the center of an image, the average width of a texture element varied between 0.98° and 1.23°. We added additional superimposed noise to the Voronoi pattern to provide dense stereo information and reduce the possibility of false correspondences. The broadband noise texture (Figure 4) was a 1600 × 1600 pixel image with random pixel intensities chosen from an approximately normal distribution, truncated at ±2 standard deviations. The noise pattern was scaled so that the number of pixels along the mid-line of the projected image was between 2200 and 2800. This noise texture provides rich stereo information when viewed binocularly, but does not provide an effective monocular slant cue.

Surfaces were slanted around a horizontal axis (i.e., receding in the vertical direction) by 12 different amounts: 5°, 10°, 15°, . . . , 60°. An exception was surfaces with broadband noise textures in monocular conditions. These stimuli were not expected to produce systematic perception of surface slant, so we only tested four slants: 15°, 30°, 45°, 60°. In cue conflict conditions, the slant specified by stereo had the same range of values (5°–60° or 15°–60°), but the slant specified by texture differed by ±5°.

Cue conflict stimuli constructed in the same way as previous studies (Knill & Saunders, 2003; Hillis et al., 2004) A distorted texture was first generated by projecting the vertices of the texture for a cyclopean view of the surface with texture specified slant and back-projecting onto a surface with stereo specified slant. A planar surface was then rendered with the distorted texture at the stereo specified slant. An example of a cue conflict stimulus is shown in Figure 3.

Procedure
The task of an observer was to align the palm of their right hand with a slanted surface. Observers wore
a flat board with 3D markers on their right hand. A small fixation cross was first presented for 2 s, followed by the simulated surface. The appearance of the surface was the cue for response. Observers lifted their hand from the table and oriented it to match the perceived orientation of the surface, and then pushed a button to indicate that the alignment was complete. Trials were self-paced, and response times were typically 1–2 s. Observers were asked to keep their hand in a general region in front and to the right of their body, to be near the tracking apparatus, but otherwise the position of the hand and arm were not constrained. Rest breaks were provided every 5 min to avoid fatigue.

The experiment was conducted over two sessions on separate days. Each session consisted of a block of binocular trials followed by a block of monocular trials. The binocular trials were run first to help observers learn the task. These stimuli produced a more vivid percept of a slanted surface and slant estimates were close to veridical, which could serve to establish response scaling for the subsequent monocular stimuli. The binocular and monocular blocks consisted of 432 trials and 96 trials, respectively, and all conditions were randomly intermixed within blocks.

Results
Perceptual bias

The top graphs in Figure 5 plot mean orientation of the hand as a function of simulated surface slant, averaged across observers, for binocular stimuli with consistent stereo and texture information. In these conditions, slant estimates were approximately linear functions of slant, and mean estimates were close to veridical for both the noise and Voronoi textures. We performed robust regression fits of the psychometric functions to minimize the effect of outliers, using iterative reweighted least-squares and a Talwar weighting function (Holland & Welsch, 1977). For the noise texture, the linear fits for individual observers’ results have average slope of 0.68 ± 0.19 SD, and for the Voronoi texture, the slopes of linear fits averaged 0.79 ± 0.12 SD. A slope of one would indicate veridical performance. For both types of texture, the slopes were significantly less than one, noise: t(10) = 5.44, p < 0.001; Voronoi: t(10) = 5.63, p < 0.001, indicating that there was some compression of slant estimates even in binocular conditions. The slopes for the two binocular conditions were significantly different, t(10) = 3.487, p = 0.006, with less overall compression of slant estimates.
in the binocular condition with Voronoi texture than the binocular condition with noise texture. The bottom graphs in Figure 5 plot mean slant estimates as a function of simulated slant for monocular stimuli with the two types of surface texture. For the monocular condition with noise texture, slant estimates showed relatively little variation across different simulated slants. There was no significant correlation between simulated slant and slant estimates for any of the observers. This indicates a large compression of responses toward a constant slant, confirming that the noise texture, by itself, was a poor monocular cue to surface slant. For the Voronoi texture, slant estimates varied as a function of simulated slant in a nonlinear manner. At low slants, 5°–30°, slant estimates were highly compressed toward frontal and there was little modulation by simulated slant. At higher slants, estimates were closer to veridical and showed greater modulation.

The psychometric functions of individual observers showed some constant biases that can be attributed to the hand alignment task. Frontal surfaces would be expected to appear frontal in all conditions, so any systematic error in the hand orientation for a surface with zero slant would likely be due to visual-motor calibration. We computed the S = 0 intersection point of the psychometric functions from each observer and condition by performing linear fits to the responses at slants from 5° to 30°. These intercepts represent the expected hand orientations for a frontal surface in each condition. Figure 6 shows the intersection points for individual observers. The three graphs plot the intersection points for binocular noise texture, monocular Voronoi texture, and monocular noise texture conditions as a function of the intersection points in the binocular Voronoi texture condition. Observers showed considerable individual differences in the intersection points (SD = 12.45°–17.23°), which were highly correlated across conditions (r = 0.796–0.982). This suggests that observers had idiosyncratic biases in the hand position associated with a frontal surface, which were generally consistent across conditions. Such constant biases would not be surprising because observers received no feedback about the accuracy of their estimates. We estimated the perceptual gain as a function of slant for each observer and condition using the difference between the average hand orientation for a surface with slant S and the expected hand orientation for a frontal surface: $g(S) = (H(S) - H(0)) / S$. This removes the influence of any constant bias in the mapping from perceived slant to hand orientation.

The left panel of Figure 7 plots mean perceptual gain averaged across observers as a function of simulated slant. In the monocular condition with Voronoi texture, perceptual gain was highly dependent on slant, as evidenced by a significant linear trend, $t(10) = 7.02$, $p < 0.001$. If perceptual gain is interpreted as the weighting of slant cues relative to a frontal tendency, this result indicates that the weighting of texture information increased with slant, as would be expected based on the reliability of texture information. For the binocular conditions, the perceptual gain was approximately constant across different stimulated slants, with no significant linear effects, binocular noise: $t(10) = 0.62, p = 0.55$, binocular Voronoi: $t(10) = 1.42, p = 0.19$. There was no detectable difference between the overall perceptual gain in the two binocular conditions when averaged across slants, $t(10) = 0.58, p = 0.58$. However, if the perceptual gain depends on reliability of slant information, a difference between the binocular conditions would only be expected at high slants, when Voronoi texture provides reliable information. To test for this interaction, we subdivided the binocular conditions into high slant (35°–60°) and low slant (5°–30°).
Figure 7. Perceptual gain as a function of slant in Experiment 1 for the monocular condition with Voronoi texture (red), the binocular condition with noise texture (blue), and the binocular condition with Voronoi texture (black). The left graph shows perceptual gain computed relative to frontal: \( g(S) = (H(S) - H(0)) / S \). The other graphs show perceptual gain computed relative to a nonzero reference slant, \( g(S) = (H(S) - H(S_{\text{ref}})) / (S - S_{\text{ref}}) \), where \( S_{\text{ref}} = 10^\circ \) (middle) or \( S_{\text{ref}} = 20^\circ \) (right). Error bars depict standard errors.

30\(^\circ\) subsets and did an ANOVA on the average perceptual gain in these subsets. We found a significant interaction between slant (low vs. high) and texture (noise vs. Voronoi) for the binocular conditions, \( F(1, 10) = 7.6, p = 0.020 \). When we compared the binocular conditions separately at high and low slants, we found a significant difference at high slants, \( t(10) = 2.55, p = 0.029 \), but not at low slants, \( t(10) = 0.5, p = 0.65 \). The difference between the binocular conditions is consistent with greater weighting of slant cues relative to a frontal tendency when both stereo and texture information are available (binocular Voronoi) than when only stereo information is available (binocular noise). This difference was observed at high slants, when texture information is reliable, but not at low slants.

Although the expected bias in slant perception would be toward frontal, the results suggest that the overall bias may have been toward a slant greater than zero. If the bias were toward frontal, then one would expect the psychometric functions to converge for surfaces with slant of zero. However, the intercepts were higher on average in the monocular conditions (13.9\(^\circ\) and 22.3\(^\circ\)) than the binocular conditions (8.7\(^\circ\) and 11.8\(^\circ\)). Statistical comparisons of the intercepts found significant differences between the binocular Voronoi and the monocular Voronoi conditions, \( t(10) = 3.13, p = 0.011 \), the binocular Voronoi and monocular noise conditions, \( t(10) = 5.5, p < 0.001 \), the binocular noise and the monocular noise conditions, \( t(10) = 3.35, p = 0.007 \), and the monocular Voronoi and the monocular noise conditions, \( t(10) = 3.16, p = 0.010 \). These differences indicate that the psychometric functions did not converge for surfaces with zero slant. Inspection of the mean psychometric functions (Figure 5) suggests that they instead converged for surfaces with slant of around 10\(^\circ\).

We recomputed perceptual gain as a function of slant using alternative reference slants to test how this would change the results. Perceptual gain was computed as \( g(S) = (H(S) - H(S_{\text{ref}})) / (S - S_{\text{ref}}) \), where \( S_{\text{ref}} = 10^\circ \) or 20\(^\circ\). This would be the appropriate measure of perceptual gain for a model in which slant information from texture and/or stereo is combined with an overall bias toward a slant of \( S_{\text{ref}} \) rather than toward a slant of zero. We excluded cases where the difference \( (S - S_{\text{ref}}) \) was less than 10\(^\circ\) because small differences in slant lead to unstable ratios. The results are shown in middle and right panels of Figure 7. The pattern of results is largely the same as before. For all reference slants, perceptual gain in the monocular condition with Voronoi texture shows a large increase with slant, while perceptual gain in the binocular conditions is approximately constant across slants and larger for the binocular condition with Voronoi texture. The main consequence of using a higher reference slant is that the perceptual gain in the monocular Voronoi condition is larger compared to the perceptual gain in the binocular noise condition at high slants. When perceptual gain was computed relative to \( S_{\text{ref}} = 0^\circ \) (left panel), the perceptual gain in the monocular Voronoi condition remains smaller than the perceptual gain in the binocular noise condition even at the highest slant, while when perceptual gain was computed relative to \( S_{\text{ref}} = 20^\circ \) (right panel), the perceptual gain in the monocular Voronoi condition becomes larger at the highest slant. Thus, if the perceptual bias was actually toward a nonzero slant rather than toward frontal, then the perceptual gain computed relative to \( S_{\text{ref}} = 0^\circ \) would underestimate the influence of texture information on slant perception relative to the influence of stereo information.

**Cue weights in conflict conditions**

Figure 8 plots mean slant estimates from cue conflict conditions. The left graph shows results for the noise
texture and the right graph shows results for the Voronoi texture. The two sets of points on each graph correspond to conditions where the texture slant is 5° larger than the stereo slant (blue) or 5° lower than stereo slant (red). For the noise texture, the positive and negative conflict conditions produced equivalent results, indicating little or no influence of this texture on perceived slant. For the Voronoi texture, one can see a divergence in the mean slant estimates from positive and negative conflict conditions, indicating an influence of texture information on perceived slant.

The relative weighting of texture and stereo information was estimated at each base slant by comparing conditions with consistent and conflicting cues. We used the cue conflict results for each base slant to compute a slope $\beta_{\text{tex}}$ that represents the effect of changing just the texture cue. For each slant $S$, we fit a line to the median slant estimates in conditions with stereo slant of $S$ and texture slant of $S - 5$, $S$, and $S + 5$, and the slope of this line was $\beta_{\text{tex}}$. We then computed another slope $\beta_{\text{st}+\text{tex}}$ representing the effect of changing both texture and stereo cues. This would be the local slope of the psychometric function from binocular conditions with consistent Voronoi texture (Figure 5). Because the psychometric functions of individual observers were approximately linear ($R^2$: 0.59 – 0.92), we used the slope of the overall function as an estimate of $\beta_{\text{st}+\text{tex}}$ for all base slants. The ratio $\beta_{\text{tex}}/\beta_{\text{st}+\text{tex}}$ was the measure of the texture cue weight. This weight indicates the effect of changing texture as a proportion of the effect of changing both texture and stereo.

Figure 9 plots texture weights as a function of slant for the two texture types. For the noise texture, the texture weights were not significantly different from zero at any slant ($p = 0.096–0.926$), indicating that the texture gradient from the noise texture had little or no effect on slant estimates. For the Voronoi texture, significant effects of texture were observed at slants of 30°–50° ($p < 0.013$), 20° ($p = 0.042$), and 60° ($p = 0.002$). The average texture weights at higher slants, 35°–60°, was greater than the average texture weights at low slants, 5°–30°, $t(10) = 4.49$, $p = 0.001$, indicating an overall increase in texture weight with simulated slant. However, the texture weights did not show a simple monotonic pattern. At the highest slants, there was an unexpected nonlinearity: average texture weights at slants of 35°–45° were higher than at slants of 50°–60°, $t(10) = 3.05$, $p = 0.012$.

**Effect of adding texture information**

A qualitative prediction of the Bayesian model, discussed in the Introduction, is that perceived slant from combined stereo and texture information could be higher than perceived slant from stereo alone even when perceived slant from texture alone is less than the perceived slant from stereo alone. In the consistent cue conditions, we did not observe a significant increase in slant estimates when texture was added to stereo information. Over the range of slant where the Voronoi texture was found to influence perceived slant, 30°–50°, there was no overall difference in the mean slant estimates for the binocular condition with noise texture and Voronoi texture, $F(1, 10) = 0.25$, $p = 0.63$. Adding texture information did not decrease slant estimates toward the biased estimates in the monocular condition, but it also did not produce a detectable increase in slant estimates. However, this null finding could be due to the fact that slant estimates were close to veridical even without texture information. In this situation, the expected increase in perceived slant due to texture information would be small and potentially difficult to detect.

Figure 10 shows another set of comparisons that demonstrates the predicted qualitative effect. Mean slant estimates from binocular conditions with noise texture (black squares) are plotted together with slant estimates in binocular conditions with Voronoi texture.
and a $+5^\circ$ cue conflict (blue triangles) and slant estimates in monocular conditions with the same texture as the cue conflict conditions (red circles). Even though the texture information specifies a slant that is $5^\circ$ larger than stereo information, slant estimates in the monocular condition were lower on average than in the stereo-only condition, $F(1, 10) = 5.28, p = 0.044$, due to the larger amount of perceptual bias in the monocular condition. However, slant estimates in the condition with combined stereo and texture information were higher than from stereo information alone, $F(1, 10) = 6.42, p = 0.030$. In these cases, adding texture information shifted slant estimates toward the veridical slant specified by texture rather than toward the biased slant estimates in monocular conditions.

**Discussion**

**Cue weights from slant estimation**

Using our hand-alignment task, observers were able to make reasonably accurate estimates of 3D surface slant when rich information was available. In binocular conditions, slant estimates were approximately linear functions of veridical slant, with means slopes of 0.79 for the Voronoi texture and 0.68 for the noise texture. Some previous studies have used a similar hand-alignment task and also observed that slant estimates were close to veridical (Norman et al., 2009; Durgin et al., 2010). While we observed some bias and compression of range even in the richest condition, slant judgments were systematic and approximately linear functions of veridical slants.

The relative weighting of texture and stereo cues was estimated from responses in cue conflict conditions. We observed a significant influence of conflicting texture information at slants of $35^\circ$–$55^\circ$. At these slants, texture information was not simply vetoed by information from stereo. This is contrary to the results of the cue conflict experiment by Todd et al. (2010), which showed no detectable influence of texture in binocular conditions. This discrepancy may be due to the larger cue conflicts tested by Todd et al., as discussed previously.

If cue integration were optimal, one would expect texture weights to systematically increase with slant. We observed larger texture weights at high slants than at low slants, consistent with this prediction. However, we also observed an unexpected nonlinearity: the texture weight at the slant of $55^\circ$–$60^\circ$ was significantly lower than at $45^\circ$–$50^\circ$. Experiment 2 tested possible explanations for this nonlinearity.

![Figure 9. Texture cue weight as a function of simulated slant in Experiment 1 for the noise texture (left) and Voronoi texture (right). Error bars depict standard errors of the means.](image)

![Figure 10. Mean slant estimates from cue conflict conditions with stereo slants of $30^\circ$–$50^\circ$ and texture slants of $35^\circ$–$55^\circ$ (blue triangles) plotted together with slant estimates from stereo only conditions (black squares) and texture only conditions (red circles) with the same slants as the cue conflict conditions. Although the slant specified by texture is $5^\circ$ higher than the slant specified by stereo, the slant estimates in the texture only condition were lower overall than in the stereo only condition. However, the slant estimates from combined cues were higher overall than in the stereo only condition.](image)
**Perceptual underestimation of slant from texture**

Slant estimates in the monocular condition with Voronoi texture were a nonlinear function of slants, as has been observed previously (Todd et al., 2005; Norman et al., 2009; Todd et al., 2010). At low slants, slant judgments were highly compressed toward frontal, while at higher slants the slant judgments were proportionally closer to veridical. As discussed in the Introduction, this nonlinear pattern could be explained in two qualitatively different ways: Texture information might provide biased internal estimates of slant or might provide unbiased but weak estimates that do not sufficiently counteract frontal cues.

Results from binocular conditions suggest that biased judgments in monocular conditions are not simply due to biased interpretation of the texture information. The binocular condition with noise texture provides a baseline measure of slant perception from stereo information. While the noise texture potentially provides texture information about slant, the responses to cue conflicts show that the noise texture had little if any influence on judgments. Thus, the binocular condition with noise texture was essentially a stereo-only condition, which can be compared to the binocular condition with Voronoi texture that provides both stereo and texture information. If the visual system’s interpretation of texture information was highly biased toward frontal, then any influence of texture information would be expected to reduce perceived slant. However, no such influence was observed in our results. At slants of 20°–60°, slant judgments from texture only (i.e., monocular Voronoi) were lower than slant judgments from stereo-only (binocular noise), yet slant judgments from combined cues (binocular Voronoi) were equal or greater than slant judgments in the stereo-only condition. Even though slant judgments in the monocular conditions were highly biased toward frontal, texture information in the combined cue condition did not appear to reduce perceived slant.

Texture information was not simply ignored in binocular conditions, which could otherwise explain why there was no increase in frontal bias. The results from the cue conflict conditions show that the Voronoi texture did have an influence on perceived slant even when stereo information was available. However, the influence of texture was toward veridical slant rather than toward the mean slant estimates from monocular conditions. In some cases, slant estimates were higher in conditions with combined stereo and texture information than from either cue alone (Figure 10). These results are not consistent with the biased texture cue explanation of Todd et al. (2010). Texture information influenced perceived slant in binocular conditions, as evidenced by results from conflict conditions, but this influence did not lead to increased underestimation of slant.

The results are generally consistent with the Bayesian model described earlier, in which stereo and texture information is integrated with frontal cues or a prior. Perceptual compression varied across conditions in a manner consistent with the reliability of slant information. In the monocular conditions with Voronoi texture, there was less proportional bias at high slants than at low slants, and the combined cue condition showed less perceptual compression than the stereo-only condition.

**Response bias**

The hand alignment task provides an indirect measure of perceived slant that would likely introduce some responses bias. We found that observers made errors in matching the orientation of their hand to a frontal surface, which are likely due to response bias. Norman et al. (2009) observed similar constant biases when slant was estimated with a manual task, but did not observe constant biases in verbal slant estimates. In addition to constant biases, there may have been some bias in scaling of responses. The estimates of cue weights (Figure 9) would be invariant both constant bias and multiplicative scaling. However, the measures of perceptual compression toward frontal (Figure 7) would be sensitive to response bias. Any multiplicative scaling in the mapping from perceived slant to hand orientation would directly scale the measures of perceptual gain. If the response bias were constant across conditions, the effect would be an overall scaling of perceptual gain, which would not change the qualitative results. If response bias were not constant, then this could also affect the relative perceptual gain across different conditions.

Our results suggest that there may have been some difference in response bias for the monocular and binocular conditions. Frontal surfaces would be expected to appear frontal in all conditions, yet the average orientation of the hand for a frontal surface was slightly higher in the monocular conditions than in the binocular conditions. The psychometric functions appeared to converge for surfaces with slant of around 10° rather than for frontal surfaces. One possible interpretation of these results is that slant information from texture and stereo is combined with an overall perceptual bias toward a slant of 10°, rather than a perceptual bias toward frontal. However, this is inconsistent with the subjective appearance of the stimuli, which appear to be frontal surfaces when simulated slant is zero, and with previous results that reported a bias toward frontal (Norman et al., 2009; Todd et al., 2010). An alternate interpretation is that there were differences in response bias across conditions. Suppose that when slant information is ambig-
uous, like monocular conditions with low slant, observers have some additional response bias toward the average orientation of the hand during the experiment. This could potentially explain why the intercepts of the psychometric functions in the monocular conditions are slightly higher than in the binocular conditions, even if the perceptual bias was toward a frontal surface.

The relative scaling of responses in our binocular and monocular conditions can be assessed indirectly by comparison to results from Todd et al. (2010). In one experiment, observers adjusted the slant of a surface specified by binocular noise to match the perceived slant of a surface specified by monocular texture, for stimuli with a similar field-of-view as tested here. Todd et al. found that a texture-only stimuli with 50° slant was matched to a stereo-only stimuli with 30° slant. Our results indicate that the average slant estimate of a texture-only stimuli with 50° slant would be approximately the same as the average slant estimate from a stereo-only stimuli with 36° slant, which is similar to the matching results of Todd et al. (2010). This general agreement suggests there was not a large difference in response biases for monocular and binocular conditions in our experiment.

**Experiment 2**

The goals of Experiment 2 were: (a) to replicate the main results of the previous experiment and (b) to investigate why texture cue weights decreased at the highest slants in Experiment 1. One possibility is that the ±5° cue conflicts were too large for normal integration in conditions with high slant, for which monocular discrimination thresholds can be as low as 3° (Knill & Saunders, 2003). Experiment 2 tested smaller cue conflicts of ±2° as well as ±5°. Another possibility is that the reduced texture weights at high slants were due to an attenuation of range effect. Experiment 2 also tested a wider range of slants in consistent cue conditions, 5°–70°, which fully encompassed the range of responses expected in cue conflict conditions.

**Methods**

**Participants**

Eleven adults (three males and eight females) at University of Hong Kong were recruited to participate in Experiment 2. One female participant was excluded because she reported that she had some problems in positioning her hand during the experiment and her performance showed abnormal negative bias. The data reported are from the other 10 participants. All the participants had normal or corrected-to-normal visual acuity, and passed a screening test for stereo acuity. All participants were naive as to the purpose of the study and were paid for participating. The procedures were approved by and conform to the standards of the Human Research Ethics Committee for Non-Clinical Faculties.

**Apparatus and stimuli**

The apparatus was the same as in Experiment 1. The stimuli were also the same except for some changes in slant and cue conflict conditions. For consistent cue stimuli, stimulated slants were 5°, 10°, … 70°. We did not include monocular conditions with noise texture in Experiment 2, because the results of Experiment 1 showed that the noise texture was not an effective monocular slant cue. Cue conflict conditions were tested around a subset of base slants: 20°, 30°, 40°, 50°, or 60°. For each of these base slants, two magnitudes of cue conflict were tested: ±2° and ±5°.

**Procedure**

The slant judgment task was the same as in Experiment 1 except for one modification. Between trials, observers were asked to move their hand back to a resting position on the table, rather than leaving their hand in the position of their last judgment. This reduced fatigue and ensured that the initial position of the hand was constant. The procedure was otherwise the same as in Experiment 1. Observers performed two experimental sessions, each consisting of one binocular session followed by a monocular session. Slant and conflict conditions were fully randomized within blocks. Trials were self-paced, and rest breaks were provided every 7.5 min.

**Results and discussion**

Figure 11 plots mean slant judgments from Experiment 2 as a function of slant for binocular conditions with consistent slant cues and the monocular condition with Voronoi texture. The general results were the same as in the previous experiment. Slant judgments in the binocular conditions were approximately linear functions of slant, with slopes of 0.60 ± 0.13 SD for the noise texture and 0.66 ± 0.14 SD for the Voronoi texture. These slopes were significantly lower than one for both binocular conditions, binocular noise: t(9) = 8.60, p < 0.001; binocular Voronoi: t(9) = 9.33, p < 0.001. As in the previous experiment, the slopes in binocular Voronoi condition were significantly higher than in the binocular noise condition, t(9) = 3.10, p = 0.013, indicating less overall compression of slant estimates when texture information was available.
the monocular condition, slant judgments again showed the same nonlinear pattern as in the previous experiment, with proportionally greater compression toward frontal at low slants than at high slants.

The psychometric functions again showed constant biases that varied considerably across observers (SD = 9.8–10.5), and were highly correlated across conditions (r = 0.87–0.91). However, unlike in the previous experiment, there were no significant differences between the intercepts in the three conditions, binocular Voronoi vs binocular noise: t(9) = 0.92; p = 0.38; binocular Voronoi versus monocular Voronoi: t(9) = 0.32, p = 0.78; binocular noise versus monocular Voronoi: t(9) = 0.54, p = 0.60. The psychometric functions from different conditions appeared to converge for frontal surfaces, so there was no evidence for any difference in response bias.

Figure 12 plots perceptual gain as a function of slant for the monocular condition with Voronoi texture (red), the binocular condition with noise texture (blue), and the binocular condition with Voronoi texture (black). Error bars depict standard errors.

Figure 11. Slant estimates as a function of simulated slant in Experiment 2 for binocular conditions with consistent cues (left, middle) or monocular conditions (right). The graphs plot mean slant estimates averaged across observers, and error bars depict standard errors.

Figure 13 plots results from cue conflict conditions in Experiment 2. The graph plots mean slant judgments as a function of the slant conflict, with the separate lines indicating results from conditions with different base slants. At slants of 20° and 30°, varying the slant specified by texture had no significant effect on judgments, 20°: F(1, 9) = 0.035; 30°: F(1, 9) = 4.4, p = 0.066. At higher slants, slant judgments were affected by conflicting texture information, as evidenced by significant linear trend, 40°: F(1, 9) = 7.649, p = 0.022; 50°: F(1, 9) = 10.026 p = 0.011; 60°: F(1, 9) = 13.184, p = 0.005. To evaluate whether this effect was linear, we tested for cubic trends. There was no overall cubic trend, F(1, 9) = 0.181, nor cubic trends at any of the
base slants when tested separately ($p > 0.157$). Even at the highest base slant, the effect of conflicting texture information was approximately linear.

Texture cue weights were computed at each base slant based on the linear effect of conflicting texture information (i.e., slopes in Figure 13) relative to the effect of changing both texture and stereo slant (i.e., the overall slopes in Figure 11). The resulting texture weights are plotted in Figure 14. Texture weights were significantly greater than zero at slants of $30^\circ$ – $60^\circ$ ($p, 0.029$) but not at $20^\circ$ slant, $t(9) = 0.016, p = 0.929$. To test for an overall increase in texture weights with slant, we combined results from the lowest and highest slant conditions to reduce noise. The average texture weight at slants of $50^\circ$ – $60^\circ$ was significantly higher than the average texture weight at slants of $20^\circ$ – $30^\circ$, $t(9) = 3.32, p = 0.009$. Unlike the previous experiment, we did not observe a reduction in texture weight at the highest slant, $50^\circ$ versus $60^\circ$: $t(9) = 0.655, p = 0.529$. The nonlinearity in cue weights observed in Experiment 1 might have been due to a range effect, or noise in estimates of cue weights.

As in Experiment 1, the results of Experiment 2 were largely consistent with a Bayesian model that integrates stereo and texture information with a frontal tendency. The results of cue conflict conditions indicate that texture information contributed to perceived slant even in binocular conditions, yet the influence of texture did not lead to greater underestimation of slant. Cue weights and perceptual bias varied with slant in a manner consistent with the reliability of texture information. In binocular conditions, texture had a larger influence relative to stereo at high slants, and in monocular conditions, there was less perceptual underestimation at high slants. One result from Experiment 1 that was inconsistent with the Bayesian model—the reduction in cue weights at the highest slant tested—was not observed in Experiment 2.

Experiment 2 varied the amount of slant conflict to test whether the effect of conflicting texture information was approximately linear. We found that the effect of conflicting texture information was proportional to the amount of conflict, with no detectable deviation from linearity. This suggests that the small cue conflicts tested here were within a range that allowed cue integration rather than cue vetoing.

### General discussion

#### Perceptual underestimation of slant from texture

The main goal of this study was to test possible explanations for perceptual underestimation of slant from texture. We found that estimates of slant from texture in monocular conditions were biased toward frontal in a nonlinear manner, with proportionally more underestimation at low slants. This is consistent with some previous findings (Todd et al., 2005; Watt et al., 2005; Norman et al., 2009; Todd et al., 2010; Durgin et al., 2010). These biases might reflect inaccurate use of texture information, as suggested by Todd et al. (2010). They could also arise from the influence of frontal cues or a frontal prior, even if texture information were accurately interpreted by the visual system.

Our results provide evidence that perceptual underestimation of slant from texture is not solely due to
biased interpretation of texture information. If slant estimates in monocular conditions were direct indications of the visual system’s internal estimates of slant from texture, then any influence of texture information in combined cue conditions would be toward the biased estimates. This is contrary to our results. We found that texture information had an influence on perceived slant in the binocular conditions, as evidenced from cue conflict results, but the influence was toward the veridical slant specified by texture rather than the slant indicated by responses in the monocular texture condition.

These seemingly contradictory results could be explained by a model in which slant information from texture and stereo are integrated with frontal cues or a frontal prior. By this model, underestimation of slant from texture would result from texture information being too weak to counteract conflicting cues that specify a frontal surface, such as the lack of an accommodative gradient, or the influence of a frontal prior. Integrating frontal cues with slant information from texture would result in biased perception of slant even if internal estimates of slant from texture were unbiased. Thus, there is no contradiction between the large biases observed in monocular conditions and more accurate slant estimates when texture and stereo information are both present.

Another problem for the scaling contrast model proposed by Todd et al. (2007) and Todd et al. (2010) is the large change in perceptual gain as a function of slant that we observed in the monocular texture conditions. For a given field of view, scaling contrast across the image of a planar surface is proportional to the tangent of the surface slant. Although this relationship is nonlinear, the deviations from linearity occur primarily at high slants. For slants up to 50°, the ratio of scaling contrast to slant varies by less than 37%. If nonlinear perception of slant from texture were due to the nonlinear relationship between slant and scaling contrast, one would expect perceptual gain to be approximately constant over this range. However, we observed a large increase in perceptual gain as slant was increased from 20° to 50°. This is inconsistent with the small amount of nonlinearity in scaling contrast over this range of slants.

In contrast, a Bayesian model that incorporates frontal cues or a frontal prior would predict the large observed effect of slant on perceptual gain. The reliability of texture information increases dramatically as slant is increased from 0° to 50° (Knill, 1998a), and our results from cue conflict conditions indicate a large increase in the weighting of texture information relative to stereo information over this range. If the reliability of texture information determines the amount of bias toward frontal, then one would expect a similar increase in perceptual gain as a function of slant, consistent with our results.

While this model requires an assumption of some additional factor beyond the slant information provided by texture and stereo, this is not an implausible assumption. Previous studies have observed an influence of accommodative cues on slant perception (Watt et al., 2005; Norman et al., 2009). The artificial displays used in our experiments did not present an accommodative gradient, and therefore provided some conflicting information that the surface was frontal. The possibility of a frontal prior for slant perception is also plausible. Perceptual underestimation of depth in situations with limited information is a general phenomenon, and some researchers have proposed a general implicit bias to explain such findings (Gogel, 1965; Ooi, Wu, & He, 2006; Wu, Ooi, & He, 2004). There is evidence for the contribution of nonuniform priors in other contexts, such as perception of 2D motion (Weiss, Simoncelli, & Adelson, 2002) and perception of light source direction (O’Shea, Agrawala, & Banks, 2010). Thus, there is reason to expect an influence of frontal cues as well as a frontal prior, and either of these factors could provide the additional influence required to explain our findings.

Backus and Banks (1999) proposed a similar model to explain how perceived slant from stereo cues is modulated by viewing distance in cue conflict situations. The model assumes that stereo slant cues are optimally combined with some additional non-stereo information that specifies a frontal surface. Viewing distance reduces the reliability of horizontal and vertical disparity cues in different ways, which would affect the optimal weighting of stereo and non-stereo cues. Backus and Banks (1999) found that this could account for the effect of distance on various perceptual biases.

A Bayesian model that incorporates frontal cues and/or a frontal prior could explain some other results for perception of slant from texture. Todd et al. (2005) observed differences in perceptual bias across conditions with different texture types and field of view. The stimuli were monocular views of V-shaped textured surfaces, and observers judged the dihedral angle in depth. Todd et al. observed more underestimation of depth for stimuli with small field of view than a large field of view, and more underestimation for surfaces covered with irregular blobs or contours than for surfaces covered with regular blobs or contours. The field of view effect could potentially be explained by use of scaling contrast as a slant cue; Todd et al. (2005) demonstrate that there is a close linear relationship between scaling contrast and mean slant estimates averaged across texture types. However, use of scaling contrast does not provide a direct explanation for the differences across texture types. One would also have to
explain why variability in local measures of texture scaling leads to bias in the perceived slant of the overall surface. Todd et al. suggest that local measures might be blurred across the image due to spatial pooling, thereby reducing the overall scaling contrast. However, this seems insufficient to account for the large observed effects of texture regularity. For example, when concave surface were presented with 40° field of view, the perceptual gain for the irregular blobs was about half of the perceptual gain for regular blobs. If one assumes that there is some additional frontal influence that is integrated with texture information, then the effects of both field of view and texture regularity can be explained. Decreasing either texture regularity or field of view reduce the reliability of texture information. If integrated optimally, one would therefore expect less weighting of texture information relative to conflicting frontal cues or a frontal prior, and correspondingly more perceptual underestimation. A Bayesian model provides a natural explanation for a relationship between reliability of information and amount of perceptual bias.

The intrinsic constraint (IC) model proposed by Domini, Caudel, and Tassinari (2006) and Domini and Caudel (2009, 2010) also predicts a relationship between the reliability of depth information and the amount of perceptual underestimation of depth. The IC model posits that the visual system estimates affine depth structure rather than metric depths, and that magnitude of perceived depth is a function of the signal-to-noise ratio of combined depth cues rather than the veridical depth specified by various cues. The motivation of this model was some counterintuitive findings from studies of perceived depth from motion and stereo information. For example, Domini et al. (2006) found that perceived depth from combined motion and stereo information was higher than from either motion or stereo alone. They further observed that the quantitative relations could be predicted if perceived depth was assumed to be a direct function of the signal-to-noise ratio of depth information, as indicated by the just-noticeable-difference thresholds (JNDs) in single cue and combined cue conditions. When stereo and motion were both available, the JNDs from combined cues were lower than from either motion or stereo alone. Domini et al. (2006) found that the magnitude of perceived depth in combined cue conditions increased by an amount inversely proportional to the amount of reduction in JNDs.

For the conditions tested here, the IC model would predict that perceived slant from combined texture and stereo would be significantly greater than from either texture or stereo alone. We did observe a small difference between the amount of perceptual bias in the stereo-only and combined cue conditions, but the difference was much less than would be predicted by the IC model. For example, in a situation where two cues have equal reliability, the IC model predicts that perceived depth from combined cues would be a factor of √2 greater than perceived depth from either cue by itself. Based on the results of Knill and Saunders (2003), texture and stereo slant cues would be equally reliable (i.e., equal discrimination thresholds) for surfaces with slant of about 50° in our conditions. The IC model would therefore predict that perceived slant from combined stereo and texture information for this range would be substantially larger (41%) than perceived slant from stereo alone. However, we observed only a small difference (<10%). Combining texture and stereo information did not have an additive effect as predicted by the IC model, but rather appeared to increase the accuracy of slant judgments. Although the IC model predicts a relationship between the reliability of slant information and amount of perceptual bias, which we did observe, the quantitative predictions are not consistent with our findings.

Our experiments cannot distinguish whether a frontal influence is due to frontal cues or a frontal prior, but there is reason to think that accommodative cues would be the primary factor. First, there is previous evidence that accommodative blur can significantly affect 3D slant perception (Watt et al., 2005; Norman et al., 2009). Second, the magnitude of perceptual bias is larger than might be expected from a prior assumption alone. Our results suggest that the frontal influence is as strong as texture cues at slants up to 50°. A Bayesian prior represents the expected distribution of surface slants, and would not be expected to constrain slant estimation to this extent. In further research, we plan to distinguish these potential influences by independently varying the slant specified by accommodative blur cues.

Relative weighting of texture and stereo slant information

A second goal of the experiments was to address a criticism of previous studies that have used a discrimination paradigm to test for optimal integration of texture and stereo slant information. Todd et al. (2010) suggested that sequential discrimination could encourage comparison of 2D cues rather than 3D slant. If so, measured cue weights would not reflect the actual contribution of texture to slant perception.

We tested direct estimation of slant, using a hand alignment task, for cue conflict conditions similar to previous studies by Knill and Saunders (2003) and Hillis et al. (2004). The overall accuracy of slant estimates in binocular conditions suggests that task provided a good measure of perceived slant. Unlike Todd et al. (2010), we observed significant influence of
conflicting texture information in binocular conditions. The discrepancy between our findings and those of Todd et al. (2010) could be due to the size of cue conflicts, as discussed previously. Some other studies have used a visual-motor task to measure the relative weighting of monocular and binocular slant cues and similarly observed a significant influence of monocular cues (Greenwald, Knill, & Saunders 2005; Knill, 2005; Seydell, Knill, & Trommershauser, 2010).

We further found that the influence of texture was greater for surfaces with high slants than with low slants, consistent with the differing reliability of texture information. Our conditions were similar to those of Knill and Saunders (2003) except for the perceptual task, so the texture cue weights can be directly compared. Figure 15 replots the texture weights from Experiments 1 and 2 along with the texture weights and optimal predictions from Knill and Saunders (2003). The texture weights computed from slant estimation in our experiments appear to be noisier than texture weights computed from discrimination judgments, as might be expected. Otherwise, the results show good agreement. If neighboring slant conditions in our experiment are combined to reduce noise (e.g., average across 20°–40° or 40°–60° slants), the smoothed results are close to the previous measures and optimal predictions. The relative influence of texture and stereo cues on slant estimation, as measured here, would be therefore consistent with optimal integration.

The previous studies of Knill and Saunders (2003) and Hillis et al. (2004) measured slant discrimination thresholds as well as cue weights to test whether integration of slant cues was statistically optimal. Discrimination thresholds from single cue conditions provide measures of the reliability of difference cues, from which one can compute the ideal weights assuming optimal integration. Both studies observed evidence consistent with optimal integration: The relative influence of texture and stereo cues, measured from cue conflict conditions, was consistent with predictions based on discrimination thresholds.

Our data do not allow an analogous test of optimality because we cannot directly estimate the reliability of slant information from texture and stereo cues. The variability of slant estimates in single cue conditions depends on factors other than the uncertainty of visual cues to slant. One factor is noise due to the estimation task. Trial-to-trial variability in slant estimates was large even in binocular conditions with high slant, with standard deviations ranging from 6°–19° for different observers. This is higher than previously observed discrimination thresholds in comparable conditions (1°–4°), suggesting that the task contributes a substantial amount of noise. Another factor is the possible influence of frontal cues or a frontal prior. These influences would tend to reduce variability at the expense of systematic bias. To the extent that frontal cues contribute to perceived slant, variability in slant estimates would not be solely due to the reliability of slant information. Thus, our data do not provide direct measures of the reliability of texture and stereo information that can be used to predict optimal weighting.

Todd et al. (2010) have also criticized the use of discrimination thresholds for predicting optimal cue weighting of stereo and texture slant information. Todd et al. argue that discrimination judgments in texture-only conditions could be based on 2D image compression rather than perceived slant, and showed that similar discrimination thresholds could be observed for orthographic projections of textured surfaces that do not elicit a 3D percept of a surface slanted in depth.

Even if a 2D strategy were used for discrimination of slant from texture in monocular conditions, however, it would not preclude use of discrimination thresholds to predict optimal weighting of texture and stereo information. In a Bayesian model, the relative influence of a cue is determined by the spread of the likelihood function. In the case of slant from texture, this would depend on the similarity of images produced by surfaces with different slants. For example, suppose that texture gradient corresponding to a surface with 20° slant was presented. If it would be hard to distinguish this image from an image produced by a surface with 0° slant or 30° slant, then the likelihood function from texture information would be spread over this range. The discriminability of images determines the spread of a likelihood function, and therefore would be the appropriate measure for predicting the influence of texture information in a Bayesian model. If ability to discriminate slant from texture depends primarily on texture foreshortening, then ability to
discriminate foreshortening in 2D images would also provide a valid estimate of the spread of the likelihood function. In other contexts, ability to discriminate low-level properties has been used to construct Bayesian ideal observer models for more complex properties. Some examples include using discrimination of 2D motion for modeling of perception of heading from optic flow (Crowell & Banks, 1996; Saunders & Niehorster, 2010), using 2D motion for modeling perception of structure from motion (Hogervorst & Eagle, 1998), and using 2D orientation discrimination for modeling perception of slant from linear perspective (Saunders & Backus, 2006b). In the case of slant from texture, ability to discriminate 2D texture compression could similarly be used to model the reliability of texture information for 3D slant perception.

To estimate the perceptual weighting of cues from discrimination judgments in cue conflict conditions, it is crucial that discrimination is based on an integrated percept of 3D slant. If observers used a 2D strategy when discriminating surfaces with combined texture and stereo information in the studies by Knill and Saunders (2003) and Hillis et al. (2004), then the estimated cue weights would not be valid. However, we observed similar cue weights using a slant estimation task, suggesting that discrimination judgments in these previous studies were indeed based on perceived slant.

While our results do not provide direct measures of the reliability of slant cues, the relative reliability of stereo and texture information could be inferred from perceptual biases as described in the Introduction (Equation 4). Figure 16 plots the texture weights predicted by Equation 4 using the observed perceptual gains from Experiments 1 and 2, along with the results of Knill and Saunders (2003). The texture cue weights predicted from perceptual bias show a similar pattern of increasing with slant, but are lower overall than the results of the previous study (solid and dashed lines) and the texture weights computed from our cue conflict conditions (Figure 15). For example, the perceptual biases suggest that texture was always less reliable than stereo in our conditions (i.e., weight less than one), while the observed cue weights suggest that texture was as reliable as stereo at slants of 40° and higher.

There are a number of possible explanations for this quantitative discrepancy. Perceptual underestimation of slant from texture could be due to both the influence of a frontal tendency as well as some errors in the interpretation of texture information; these possibilities are not mutually exclusive. Another factor could be response bias. The results of Experiment 1 suggest that the combined influence of perceptual bias and response bias lead to an overall bias toward a nonzero slant. When perceptual gain was computed relative to a slant of 10° or 20° (Figure 7, middle and right panels), the perceptual gain in the texture-only condition increased relative to the perceptual gain in the stereo-only condition at high slants. These perceptual gains would predict optimal cue weights that are closer to the observed cue weights. Another possibility is that assumptions used to derive Equation 4 are not an accurate approximation of Bayesian integration. Based on how texture gradients vary as a function of slant, likelihood functions from texture would be skewed toward frontal (Knill, 1998a). This would predict an asymmetric effect when texture is combined with other cues: Cues that specify a lower slant would have comparatively more influence than cues that specify a higher slant. This might explain why there were larger biases in texture-only conditions than would be expected given the weighting of texture relative to stereo. The present results cannot distinguish these possibilities. Apart from the discrepancy in the overall amount of perceptual compression, the results were largely consistent with Bayesian model. If the weighting of texture information relative to frontal cues varies depending on the reliability of texture information, this would explain most, if not all, of the nonlinear pattern of slant estimates observed in the monocular condition.

## Conclusion

The experiments reported here investigated how texture and stereo cues are integrated to perceive 3D slant. Previous studies using a discrimination paradigm have found evidence that texture and stereo information are optimally integrated (Knill & Saunders, 2003; Hillis et al, 2004). However, Todd et al. (2010) has questioned the validity of measuring cue weights with this discrimination paradigm. We used a slant estimation task to measure the relative influence of texture.
and stereo and observed results that were consistent with previous studies.

We tested whether perceptual underestimation of slant could be explained by a Bayesian model that includes an additional influence from frontal cues and/or a frontal prior. This model predicts that perceptual bias would depend on the reliability of slant cues. We found that slant judgments from monocular texture were more accurate at high slants than low slants, consistent with difference in reliability of texture information. Although slant estimates from monocular texture were biased, combining texture and stereo information did not reduce the accuracy of slant estimates compared to conditions with only stereo information. In some cases, slant estimates from combined texture and stereo information were higher than from either cue alone. These results are consistent with a Bayesian model that optimally integrates texture and stereo slant information with frontal cues or a frontal prior. Biased perception of slant from texture would emerge from such a model even if texture information were accurately processed by the visual system.

Keywords: slant perception, cue combination, stereo, texture, Bayesian model

Acknowledgments

This work was supported by a grant from the Hong Kong Research Grants Council, GRF HKU-752010H.

Commercial relationships: none.
Corresponding author: Jeffrey A. Saunders.
Email: jsaun@hku.hk.
Address: Department of Psychology, University of Hong Kong, Hong Kong, Hong Kong SAR.

References


three-dimensional structure-from-motion arise from noise in the early visual system. *Proceedings of the Royal Society of Biological Sciences, 265*(1406), 1587–1593.


