PERGAMON

Vision Research 38 (1998) 2635–2656

Vision Research

Ideal observer perturbation analysis reveals human strategies for inferring surface orientation from texture

David C. Knill *

Department of Psychology, University of Pennsylvania, 3815 Walnut Street, Philadelphia PA 19104, USA

Received 6 March 1997; received in revised form 29 September 1997

Abstract

Optical texture patterns contain three quasi-independent cues to planar surface orientation: perspective scaling, projective foreshortening and density. The purpose of this work was to estimate the perceptual weights assigned to these texture cues for discriminating surface orientation and to measure the visual system's reliance on an isotropy assumption in interpreting foreshortening information. A novel analytical technique is introduced which takes advantage of the natural cue perturbations inherent in stochastic texture stimuli to estimate cue weights and measure the influence of an isotropy assumption. Ideal observers were derived which compute the exact information content of the different texture cues in the stimuli used in the experiments and which either did or did not rely on an assumption of surface texture isotropy. Simulations of the ideal observers using the same stimuli shown to subjects in a slant discrimination task provided trial-by-trial estimates of the natural cue perturbations which were inherent in the stimuli. By back-correlating subjects' judgements with the different ideal observer estimates, we were able to estimate both the weights given to each cue by subjects and the strength of subjects' prior assumptions of isotropy. In all of the conditions tested, we found that subjects relied primarily on the foreshortening cue. A small, but significant weight was given to scaling information and no significant weight was given to density information. In conditions in which the surface textures deviated from isotropy by random amounts from stimulus to stimulus, subject judgements correlated well with the estimates of an ideal observer which incorrectly assumed surface texture isotropy. This correlation was not complete, however, suggesting that a soft form of the isotropy constraint was used. Moreover, the correlation was significantly lower for textures containing higher-order information about surface orientation (skew of rectangular texture elements). The results of the analysis clearly implicate texture foreshortening as a primary cue for perceiving surface slant from texture and suggest that the visual system incorporates a strong, though not complete, bias to interpret surface textures as isotropic in its inference of surface slant from texture. They further suggest that local texture skew, when available in an image, contributes significantly to perceptual estimates of surface orientation.

1. Introduction

Since Gibson first analyzed texture gradients [1], it has been well-known that optical texture patterns can provide strong cues to surface shape and orientation (see Fig. 1). We have recently shown, for example, that subjects can use texture information to discriminate differences in planar surface slant as small as 1-1/2 to 3°. Others have shown that the visual system gives a significant weight to texture information for estimating surface orientation and curvature even with the presence of conflicting stereo or motion information [2–4]. These observations have made the problem of how humans infer 3D surface structure from texture one of significant theoretical interest.

Texture information is composed of several distinct cues: perspective scaling, projective foreshortening and texture density. Each of the cues depends on one or more well-specified prior constraints for its informativeness: homogeneity, the assumption that surface textures are similar (at least, statistically) everywhere over a surface, and isotropy, the assumption that surface textures have no particular orientation bias (again, in a statistical sense). Two fundamental questions about human perception of surface geometry from texture therefore arise: how does the visual system integrate the available texture cues to make inferences about surface geometry, and what prior assumptions about surface textures does it rely on to do so?

* Fax: + 1 215 8987301; e-mail:knill@cattell.psych.upenn.edu.

0042-6989/98/$19.00
PII: S0042-6989(97)00415-X

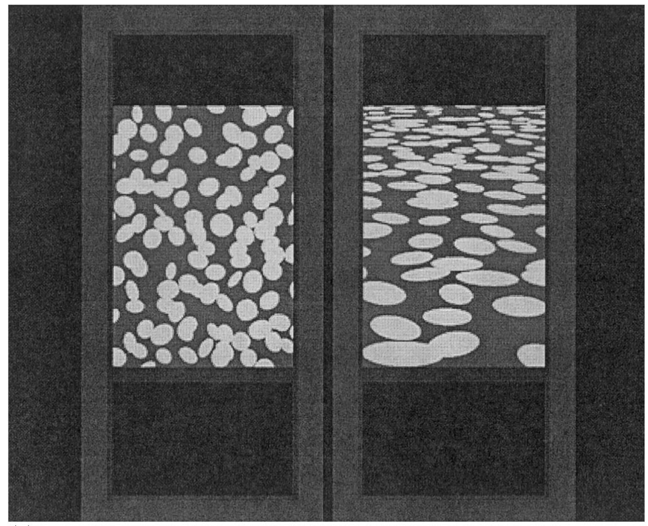
In this paper, we describe experiments and a new method of analysis designed to measure the relative contributions of the different texture cues to perceived surface slant (angle away from the fronto-parallel) and to measure the strength of a putative isotropy assumption. A number of previous researchers have investigated the same problems; however, due to the qualitative nature of that work, significant gaps in our understanding remain. Our hope is that the current work will serve to close some of these gaps as well as to introduce a powerful formal methodology which, with future applications, may further serve to flesh out our understanding of shape from texture perception.

Our approach is similar in spirit to previous work, in that it relies on a form of cue perturbation paradigm [2,3,5–8]. Rather than artificially inducing cue conflicts, however, we rely for our analysis on the natural cue perturbations inherent in images of stochastic surface textures. The recent development of ideal observers for each of the different texture cues [9–11] allows us to measure these natural perturbations, which we can in turn correlate with subjects' judgements in a slant discrimination task to estimate cue weights, test whether subjects assume isotropy and so on. Because of the important role played by the ideal observers, we will refer to the method as ideal observer perturbation analysis.

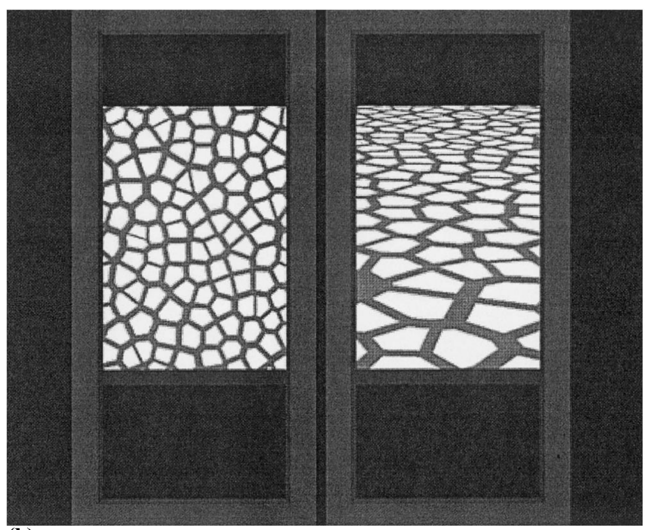
In the current paper, we describe two experiments and associated analyses designed to measure the weights given by the visual system to different texture cues and to measure the strength of the isotropy assumption imposed by the visual system. As part of our analysis, we will also consider questions of how cue weights and the strength of the isotropy assumption change across stimulus conditions. The second section of the paper provides a brief overview of the structure of texture cues. The third section introduces the logic of ideal observer perturbation analysis. The fourth section applies the analysis to the results of a simple experiment measuring subjects' abilities to discriminate surface slant (orientation away from the line of sight) from monocular images of planar, isotropic textures. The analysis provides estimates of the weights given by subjects to different texture cues for performing the discrimination task. In the fifth section, we apply the analysis to an experiment designed to measure the strength of subjects' isotropy assumptions. The final section discusses the results and relates them to previous psychophysics to draw conclusions about human processing strategies for estimating planar surface orientation from texture.

2. The structure of texture information

Locally, perspective projection distorts a texture pattern in two distinct ways: by scaling the texture and by distorting its shape. The 'size' of a local texture patch is scaled by an amount inversely proportional to the distance of the patch from the nodal point of the eye. Similarly, the shape of the patch is foreshortened in the direction of local surface tilt by an amount proportional to the cosine of the slant of the surface relative to the local line of sight. Both of these effects are first-order approximations of the local perspective distortion of texture. They vary with spatial position in a predictable way as a function of surface geometry. For



(a)



(b)

Fig. 1. Two examples of texture patterns which elicit strong percepts of planar surfaces oriented in depth. Each image contains surfaces at 0 and 65° slant. In the experiments described here, subjects viewed stimuli like these and made judgements of which surface was most slanted in depth. The patterns in (a) were created by scattering random ellipses over a surface before projection. The patterns in (b), which we will refer to throughout the text as Voronoi textures, were each created from the Voronoi diagram of a random lattice of points in the plane of the surface, a technique first introduced into shape-from-texture psychophysics by Rosenholtz and Malik (see [2, 9] for detailed descriptions of Voronoi textures).

slanted, planar surfaces, distance from the viewer changes as a simple function of position in the image. Local slant and tilt varies over an image as well, since the angle of the local line of sight changes across an image. These changes in local viewing geometry result in spatial variations in the local distortion of a texture pattern.

The scaling and shape distortion components of the local texture map (and their 'gradients') are not directly available to an observer, who only sees the effects of these distortions in the projected image. A-prior knowledge of the spatial structure of surface textures is needed to infer the form of the local distortion, and hence the local 3D shape, from image data. Since most surface textures are stochastic, the best any observer can hope for is knowledge of the statistical properties of an ensemble of surface textures. Such knowledge supports statistical 'best-guesses' about the local distortion. The two generic forms of statistical constraints which can make texture informative are homogeneity and isotropy. Homogeneity is equivalent to the statistical concept of stationarity; that the statistics of a pattern do not change with position on a surface. Isotropy refers to the lack of any orientation bias in the statistical structure of a pattern. Generally speaking, a texture ensemble is isotropic if its statistical properties can be characterized without regard to orientation.

Perspective projection induces measurable inhomogeneities and anisotropies in image texture patterns. These provide information about the orientation and shape of a surface. Foreshortening makes the distribution of texture element shape and orientation decidedly non-uniform and, in the case of isotropic textures, non-circular. Perspective scaling renders texture elements far away from the viewer, on average, smaller than those closer to the viewer. These observations lead to a natural decomposition of texture information into the so-called cues of perspective scaling and foreshortening. Both scaling and foreshortening information require an assumption of texture homogeneity to be useful. Foreshortening information, however, can take two forms with or without an assumption of isotropy. Without assuming isotropy, an observer would have to rely on spatial gradients in foreshortening to make judgements about surface orientation. With it, an observer could make orientation judgements more locally by measuring the deviation of local texture shape statistics away from isotropy.

Textures composed of discrete elements (texels) admit a third cue to surface orientation and shape density. The three texture cues can be made independent by equating scaling information with the statistical distribution of texel sizes, foreshortening information with the distribution of texel shape and orientation (defined in an appropriately scale-independent way [12]) and density with the distribution of texel positions. This latter information, for most textures, is not properly captured in a simple density measure, but rather in the relative positioning of texels. We will therefore refer to the information carried by texel positions as position information, to avoid the possibly misleading semantics of the term, density.

Since we use discrete element textures in our experiments, we will adhere to the definition of texture cues given above. More specifically, since we use textures composed of elliptical texels, we define the image equivalents of the three texture cues to be

Scaling— the spatial distribution of texel lengths.
Foreshortening— the spatial distribution of texel aspect ratios and orientations.
Position— the spatial distribution of texel positions

What we have defined as foreshortening information corresponds to the cue most often referred to in the literature as compression [6,10,13,14]. Some authors', however, have defined the compression cue differently, as the gradient in absolute compression of texture elements in the direction of surface tilt [2,3]. Compression defined this way depends on both scaling and foreshortening effects of projection (and does not admit a straightforward application of the isotropy constraint). Because of the possible confusion inherent in the use of the term, compression, we will use the term, foreshortening, to refer to the distortion in texture element shapes and orientations induced purely by projective foreshortening.

3. Ideal observer perturbation analysis

3.1. Introduction

In previous work [12], we compared subjects' thresholds for discriminating planar surface slant from texture information with the thresholds of ideal observers for each of the texture cues, using the ideal observer thresholds as objective measures of the inherent uncertainty contained in the texture cues. The comparisons allowed us to draw a number of broad, qualitative conclusions about subjects' strategies for interpreting surface slant from texture. The most significant of these conclusions was that subjects relied at least in part on foreshortening information to make their discrimination judgements.

Unfortunately, comparing threshold data between human and ideal observers, as we previously did and is typically done [10] is a very blunt tool to use in analyzing perceptual strategies. It only supports the type of broad-stroke inference described above. We should, however, be able to use the ideal observer models more directly to infer subjects' perceptual strategies. In particular, we should be able to correlate subjects' discrimination judgements on each trial with the stimulus slants derived from ideal observers for each cue, in order to gain some quantitative insight into how texture cues are combined for the perception of surface slant. This would be a form of post hoc perturbation analysis which relies only on the natural perturbations inherent in stochastic texture stimuli, as measured by the perturbations in ideal observer estimates of slant.

3.2. Natural cue perturbations and the linear model

Texture scaling, foreshortening and position information provide quasi-independent cues to the orientations of planar surfaces. The visual system must somehow combine these information sources to arrive at estimates of surface orientation. For the classes of textures we are considering, the optimal way to do this is to compute likelihood functions for each of the cues and multiply them together to obtain the full-cue likelihood function from which estimates of surface orientation may be derived. When the individual cue likelihood functions approximate Gaussian distributions, the maximum likelihood estimator for surface orientation from texture reduces to a weighted sum of the maximum likelihood estimates for surface orientation from the constituent cues contained in texture patterns. In this case, we can write the maximum likelihood estimate of slant, $\hat{\sigma}$, as

$$\hat{\sigma} = w_s\sigma_s + w_f\sigma_f + w_p\sigma_p \tag{1}$$

where w_s, w_f ,w_p are the weights given to the scaling, foreshortening and position cues, respectively, and σ_s,σ_f,σ_p are the most likely slants indicated by the different cues. The weights are determined by the relative variances of the individual cue likelihood functions.

The key observation behind the ideal observer perturbation analysis is that the orientations indicated by different texture cues for surfaces with a fixed slant vary randomly from stimulus to stimulus around that slant. A concise representation of the perturbations is as a noise process added to the true slant of the surface pictured in a stimulus,

$$\sigma_s = \sigma + N_s \tag{2}$$

$$\sigma_f = \sigma + N_f \tag{3}$$

$$\sigma_p = \sigma + N_p \tag{4}$$

where N_s, N_f and N_p are independent random processes representing the natural perturbations in surface slant suggested by the scaling, foreshortening and position cues respectively.

3.3. Estimating cue weights

One way to estimate the weights in our psychophysical model would be to fix the slant of a surface and somehow measure subjects' perceptions of slant for images of a large set of textures 'painted' on that surface. Correlating subjects' perceptual estimates of slant with the maximum likelihood estimates of slant derived from each texture cue would provide a measure of the weights of the different cues. The major problem with this approach is that any practical method for measuring perceived slant (e.g. gauge figures [8]) is likely to add significant measurement noise to the estimates of perceived slant, greatly reducing the sensitivity of the analysis. Potential non-linearities in the mapping between perceived slant and reported slant create a further complication, since they can produce distortions in the measured weights.

Rather than attempt to directly measure correlations between perturbed slants and perceived slants, we estimated the sets of weights which best predicted subjects' discriminations of surface slant from texture. While indirect, estimating weights from discrimination data has the advantage that the experimental task provides a more sensitive measure of subjects' judgements of surface slant (albeit differential slant) than do direct measures of perceived slant. We describe the basic logic of the analysis here. Details of the variants used to analyze the current experiments will be given in the sections describing the experiments.

In a discrimination task, subjects compare two texture stimuli and judge which of the two has the largest slant. In the experiments described here, we fixed the test slant to be 65° and varied the comparison slant to be greater or less than the test slant. The raw data from such an experiment consists of data pairs specifying the differences in slant between comparison and test stimuli on each trial and subjects' associated responses. We model subjects' discrimination process as being based on a decision variable, $\Delta\sigma$, which is given by

$$\Delta\sigma = w_s\Delta\sigma_s + w_f\Delta\sigma_f + w_p\Delta\sigma_p + N \tag{5}$$

where w_s is the weight given to the scaling cue, w_f is the weight given to the foreshortening cue and w_p is the weight given to the position cue. N is a noise process. $\Delta\sigma_x$ is the difference in the variance-stabilized slants of the test and comparison surface suggested by cue x. Defining subjects' response on trial i to be

$$X_i = \begin{Bmatrix} 1 \text{ comparison slant judged greater} \\ 0 \text{ test slant judged greater} \end{Bmatrix} \tag{6}$$

we can write the probability of a subject judging the comparison slant to be greater than the test slant on trial i as

$$p(X_i = 1) = \Psi(\Delta\sigma_i;\ \vec{\alpha}) \tag{7}$$

$$p(X_i = 1) = \Psi(w_s\Delta\sigma_{s_i} + w_f\Delta\sigma_{f_i} + w_p\Delta\sigma_{p_i};\ \vec{\alpha}) \tag{8}$$

where $\Psi()$ is a psychometric function and $\vec{\alpha}$ is a vector of the parameters which define the shape of the psychometric function. If $\Psi()$ is a cumulative Gaussian, for example, $\vec{\alpha}$ contains one parameter for the slope of the psychometric function. For the opposite judgement, we have

$$p(X_i = 0) = 1 - \Psi(w_s\Delta\sigma_{s_i} + w_f\Delta\sigma_{f_i} + w_p\Delta\sigma_{p_i};\ \vec{\alpha}) \qquad (9)$$

using Eqs. (8) and (9), we can compute the likelihood function for any set of n responses as

$$L(w_s,w_f,w_p,\vec{\alpha}) = \prod_i^n X_i\Psi(w_s\Delta\sigma_{s_i} + w_f\Delta\sigma_{f_i} + w_p\Delta\sigma_{p_i};\ \vec{\alpha}) + (1 - X_i)\,(1 - \Psi(w_s\Delta\sigma_{s_i} + w_f\Delta\sigma_{f_i} + w_p\Delta\sigma_{p_i};\ \vec{\alpha})) \qquad (10)$$

where the values of $\Delta\sigma_{s_i}$, $\Delta\sigma_{f_i}$ and $\Delta\sigma_{p_i}$ are the differences in comparison and test slant derived from application of the scaling, foreshortening and position ideal observers, respectively, to the stimuli for trial i. We can use Eq. (10) to derive the maximum likelihood estimate of cue weights $\{w_s,w_f,w_p\}$ for a given set of responses by integrating over the psychometric parameters in $\vec{\alpha}$.

The distinction between the current method and other perturbation methods is that we rely on the natural perturbations inherent in the random structure of stochastic texture stimuli. More traditional methods induce perturbations in stimulus levels by artificially placing cues in conflict with one another [15]. As Landy et al. have pointed out, one must take care in perturbation studies to use small perturbations. The current method automatically keeps the perturbations at a theoretical minimum.

4. Experiment 1: Estimating cue weights

We previously descibed a series of experiments on subjects' abilities to discriminate planar surface slant using texture information [11]. Experiment 4 in this paper contained conditions in which subjects' thresholds approached those of several of the ideal observers (particularly, the scaling and position ideal observers). This suggested to us that we might successfully apply the perturbation analysis described above to the raw data from the experiment to derive estimates of cue weights. In the previous presentation of results, we considered only the thresholds derived from the data. Here we re-analyze the raw data using ideal observer perturbation analysis.

4.1. Methods

4.1.1. Apparatus

Stimuli were presented on the display monitor of an SGI computer. The monitor was an SGI model TFS6705, 17 in., color display with a resolution of 1280×1024 pixels. Stimuli were generated in gray-scale on the display (to the extent that equal settings of color gun voltages generated flat spectra). Since the stimuli did not contain smooth shading variations, we did not do gamma correction. Subjects viewed the stimuli presented on the monitor monocularly through a reduction screen, with their heads placed in a chin rest and resting on a front head-rest. Subjects' non-viewing eye was covered with an eye-patch to eliminate any potential for binocular rivalry. Subjects were tested in a room painted matte black to minimize secondary reflections back onto the monitor. Finally, a matte black occluder was placed over the front of the monitor to obscure the physical screen boundaries. The monitor was calibrated using test patterns of dots viewed through a piece of metal with a square grid of holes drilled in it to ensure a square geometry.

Subjects viewed the display from a distance of 28 cm, giving a total angular extent of the display area on the screen of approximately $48 \times 40°$ of visual angle.

4.1.2. Stimuli

Stimuli consisted of images of textured surfaces subtending a width of 250 pixels and a height of 324 pixels (displayed side-by-side). At the viewing distance used of 28 cm, this resulted in images of surfaces subtending $10 \times 12.5°$ of visual angle. Stimulus texture patterns were created by randomly generating surface textures consisting of sixty elliptical elements and projecting them under perspective projection onto the computer screen. The surface textures used to generate stimuli were sampled from well-defined stochastic processes for which we had previously derived ideal observers [12]. Texel positions were sampled from a constrained positioning model (to minimize texel overlap) and texel lengths and shapes were independently sampled from different prior probability distributions. The orientations of texels was chosen from a uniform distribution around the circle, leading to textures which were globally isotropic. Four classes of textures were generated for the experiment based on a crossing of two different prior distributions of surface texel sizes and two different distributions of surface texel shapes. One each of these corresponded to highly constrained distributions and the other corresponded to broad distributions. This gave rise to four texture conditions corresponding to all possible combinations of reliable and unreliable scaling and foreshortening cues (see [11], experiment 4, for details). As we use the terms reliable and unreliable, they refer to the relative reliability between conditions, unreliable simply meaning less reliable than the 'reliable' conditions.

Stimuli were presented side by side in the experiment, with each stimulus image having its own simulated window frame. The innermost boundaries of the surface images were 70 pixels from the center of the screen (including the space taken up by the inner frame), which, for the viewing conditions used, gave a separation between inner edges of the stimuli of 6° of visual angle. For each condition in an experiment, the vertical positions of a surface's boundaries as they appeared in an image were the same for both test and target stimuli, so that boundary height in the image plane did not provide a cue to surface slant.

4.1.3. Procedure

We used a two-alternative forced choice procedure in which subjects judged which of two simultaneously presented texture images appeared to be more slanted. All conditions in an experiment were randomly interleaved, including the side of the display on which the correct stimulus appeared. The screen was blanked between trials, a period which lasted anywhere from 0.5 to 1 s, depending on the time it took to generate stimuli for the next trial. Subjects were given unlimited time to view the displays on each trial, but were explicitly instructed to make judgements based on their immediate guess as to which surface was more slanted. They were told that on some trials the choice would be clear and on others it would be more ambiguous, but to stick with their first guess regardless of how uncertain it seemed. Feedback was given in the form of a summary score every 20 trials. The feedback was used simply to make the task more palatable for subjects, as pilot studies showed subjects found the experiment with no feedback extremely unpleasant and we suffered from many drop-outs. No trial-by-trial feedback was given, in order to minimize, as much as possible, the learning of simple 2D strategies for doing the task.

Test stimuli were generated to simulate a slant of 65°. Two non-parametric staircases (four-up/one-down and one-up/four-down) were interleaved for each condition to find the 85 and 15% threshold differences in slant needed for subjects to judge a comparison stimulus to have greater slant than a test stimulus at 65°.

Before starting the main part of the experiment, subjects were run in a brief demonstration version of the experiment using textures generated from surfaces with very large differences in slant (65 and 73° for test and target stimuli respectively).

4.1.4. Subjects

Subjects were drawn from the student body at the University of Pennsylvania and were paid for their participation. Subjects had normal or corrected to normal vision and were naive to vision science.

4.2. Results

4.2.1. Discrimination thresholds

The psychometric functions measured in the experiment showed a significant negative skew, consistent with the results of other experiments showing that discrimination thresholds decrease with increasing slant [12]. This led us to model the psychometric function as a cumulative Gaussian in a variance-stabilized slant domain. We used a one-parameter, log transform as the variance stabilizing function [16]. This was given by

$$\sigma' = \log_e[1 + \beta(\sigma - \sigma_0)] \tag{11}$$

where σ_0 is the point of equality; in our case, 65°[1].The resulting psychometric model can be expressed as a two-parameter function of the un-transformed slant, where β is the variance stabilizing parameter and s is the slope parameter of the cumulative Gaussian. In order to compute discrimination thresholds, we computed maximum likelihood estimates of both β and s.

$$p(\text{correct}) = \text{erf}\,(\log_e[1 + \beta(\sigma - \sigma_0)]/s) \tag{12}$$

Fig. 2 shows the thresholds of four subjects for the four conditions used in the experiment. Also shown on the graphs are the thresholds of the ideal observers for each of the three texture cues. For our purposes here, we can treat the ideal observer thresholds as measures of the magnitudes of the slant perturbations contained within each cue. We do not consider the threshold data any further here. The reader is referred to Ref. [11] for a detailed discussion of the thresholds and their implications.

4.2.2. The analytical model

The model we used to estimate cue weights was a variant of the one presented in Section 3. We found the Gaussian noise model to be a good model for subjects' judgements when applied in a variance-stabilized slant domain. We therefore applied the linear cue combination model to surface slant expressed in the transformed domain. This required applying equation Eq. (11)to the slants indicated by each of the three texture cues before combining them in a linear sum. Fig. 3 illustrates the basic psychophysical model. The decision variable, $\Delta\sigma'$, is given by

$$\Delta\sigma' = w_s\Delta\sigma'_s + w_f\Delta\sigma'_f + w_p\Delta\sigma'_p + N \tag{13}$$

where w_s is the weight given to the scaling cue, w_f is the weight given to the foreshortening cue and w_p is the weight given to the position cue. N is a Gaussian noise process. $\Delta\sigma'_x$ is the difference in the variance-stabilized slants of the test and comparison surface suggested by cue x.

[1] The log transform is derived by assuming that the standard deviation in perceived slant changes linearly in the neighborhood of the test slant.

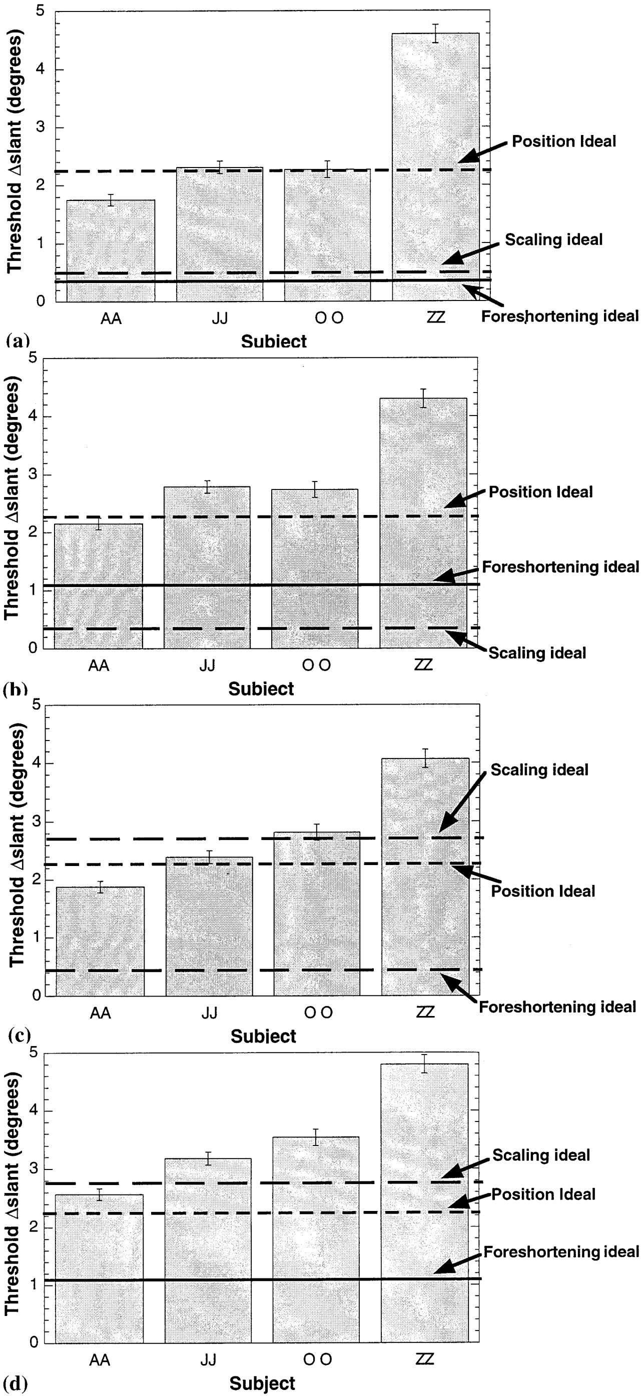


Fig. 2. Plots of subjects' 75% thresholds in each of the four experimental conditions in experiment 1: (a) reliable scaling/reliable compression; (b) reliable scaling/unreliable compression; (c) unreliable scaling/reliable compression; and (d) unreliable scaling/unreliable compression. Thresholds for the each cue's ideal observers are shown as horizontal lines in each of the plots.

In this model, the psychometric function has two free parameters besides the cue weights; the variance-stabilizing parameter, β, and the slope parameter for a cumulative Gaussian, s (see equation Eq. (12). Thus we have as the likelihood function for subjects' responses

$$
\begin{aligned}
&L(w_s,w_f,w_p,\beta,s \mid X) \\
&= \prod_i^n X_i \mathrm{erf}\left(\frac{1}{s}(w_s\Delta\sigma'_{si} + w_f\Delta\sigma'_{f_i} + w_p\Delta\sigma'_{p_i})\right) \\
&\quad + (1 - X_i)\left(1 - \mathrm{erf}\left(\frac{1}{s}(w_s\Delta\sigma'_{s_i} + w_f\Delta\sigma'_{f_i} + w_p\Delta\sigma'_{p_i})\right)\right)
\end{aligned}
\tag{14}
$$

where the values of $\Delta\sigma'_{s_i}$, $\Delta\sigma'_{f_i}$ and $\Delta\sigma'_{p_i}$ are the differences in variance-stabilized slant derived from application of scaling, foreshortening and position ideal observers, respectively, to the stimuli for trial i.

In order to estimate cue weights, we integrated the likelihood function over β and s. We further constrained the cue weights to sum to one, since the absolute magnitude of the weights is indeterminate from discrimination data. A simplex search [17] was used to find the maximum likelihood estimates of cue weights for each condition in the experiment.

4.2.3. Results of the analysis

The space of allowable weights lies in a triangular planar region in the three-dimensional space of cue weights, as illustrated in Fig. 4. The vertices of the triangle correspond to weight configurations in which only one cue is used to perform the task. The sides correspond to weight configurations in which one or another cue is unused. Fig. 5 shows an example of a likelihood function derived from one subjects' data, plotted in this triangular space. The most salient feature of the function is that it is concentrated on the side of the triangle corresponding to a position weight of zero. We found similar results for all subjects and conditions in the experiment. The apparently insignificant weight given to position information in the experiment allows us to present the results in a much more compact form as a function only of the relative weight given to scaling and foreshortening information.

Fig. 6 shows results of the perturbation analysis applied to the three conditions of the experiment in which at least one of the scaling or foreshortening cues had some non-trivial uncertainty (in one condition, both were too reliable to generate significant perturbations). The results are plotted as the likelihood functions for the foreshortening weight (the scaling weight is one minus the foreshortening weight), given subjects' responses. Table 1 summarizes the results, listing the maximum likelihood estimates of the foreshortening weight, with 95% confidence intervals.

The data has three notable features. First, as already mentioned, the position cue receives little or no weight in subjects' discrimination strategies. Second, foreshortening information is consistently weighted more strongly than scaling information. Only in the condition that scaling information is very reliable and fore-

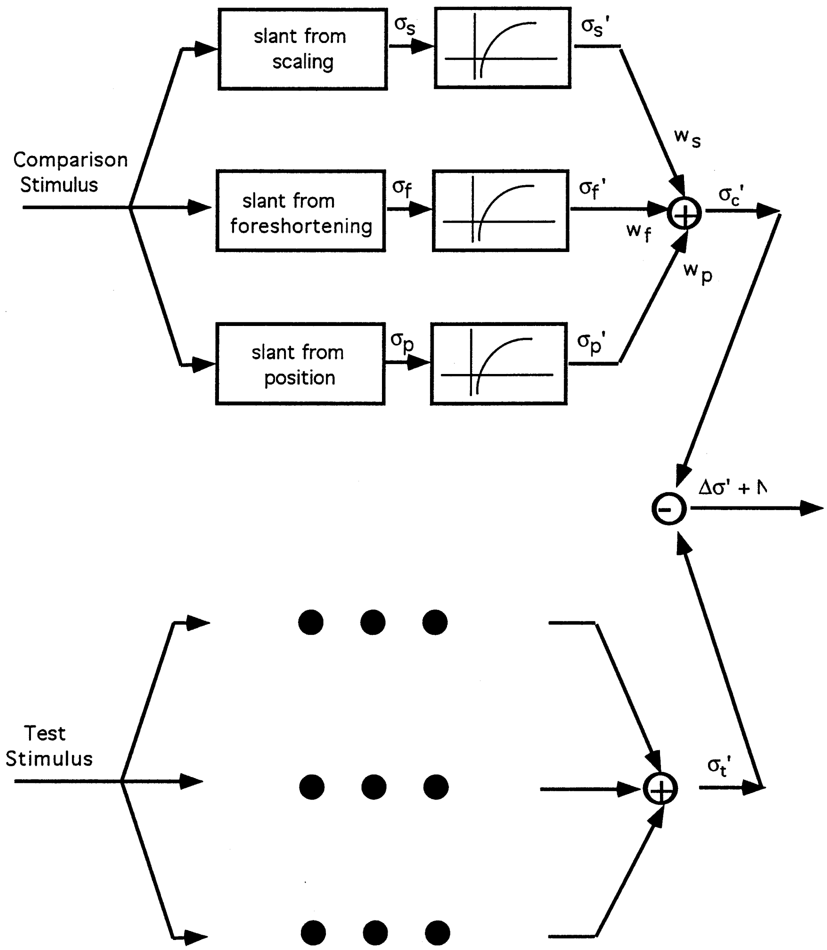


Fig. 3. The signal detection model we used for our analysis (see text for description).

shortening information is weaker is there any hint that scaling is given close to the same weight as foreshortening information. The confidence intervals on the weights estimated for this condition, however, are quite broad. Finally, there is a weak suggestion that cue weights vary consistently with relative cue reliability. For three of the four subjects, foreshortening information gets the smallest weight when scaling information is a very reliable indicator of slant and foreshortening is not. Foreshortening receives the largest weight when it is a reliable cue and scaling is not.

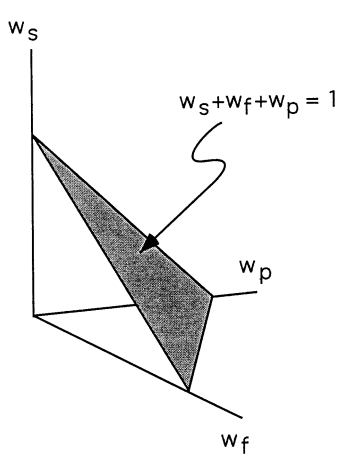


Fig. 4. The space of possible cue weights lies on a plane defined by the equation, $w_s + w_f + w_p = 1$, and is bounded by the constraint that all weights are positive.

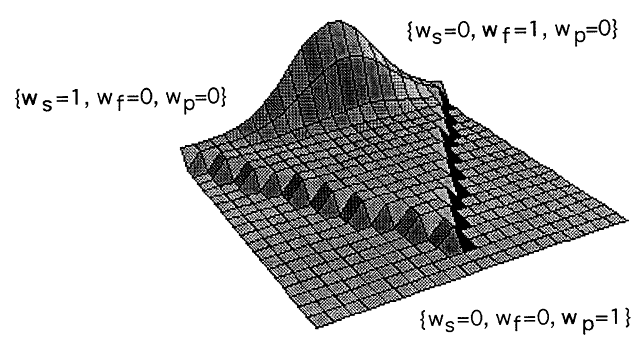


Fig. 5. The likelihood function for cue weights for one subject (subject AA) in the scaling reliable/foreshortening unreliable condition. The likelihood function is concentrated on the side of the triangle corresponding to a zero weight for the position cue. This effect is found for all subjects in all conditions.

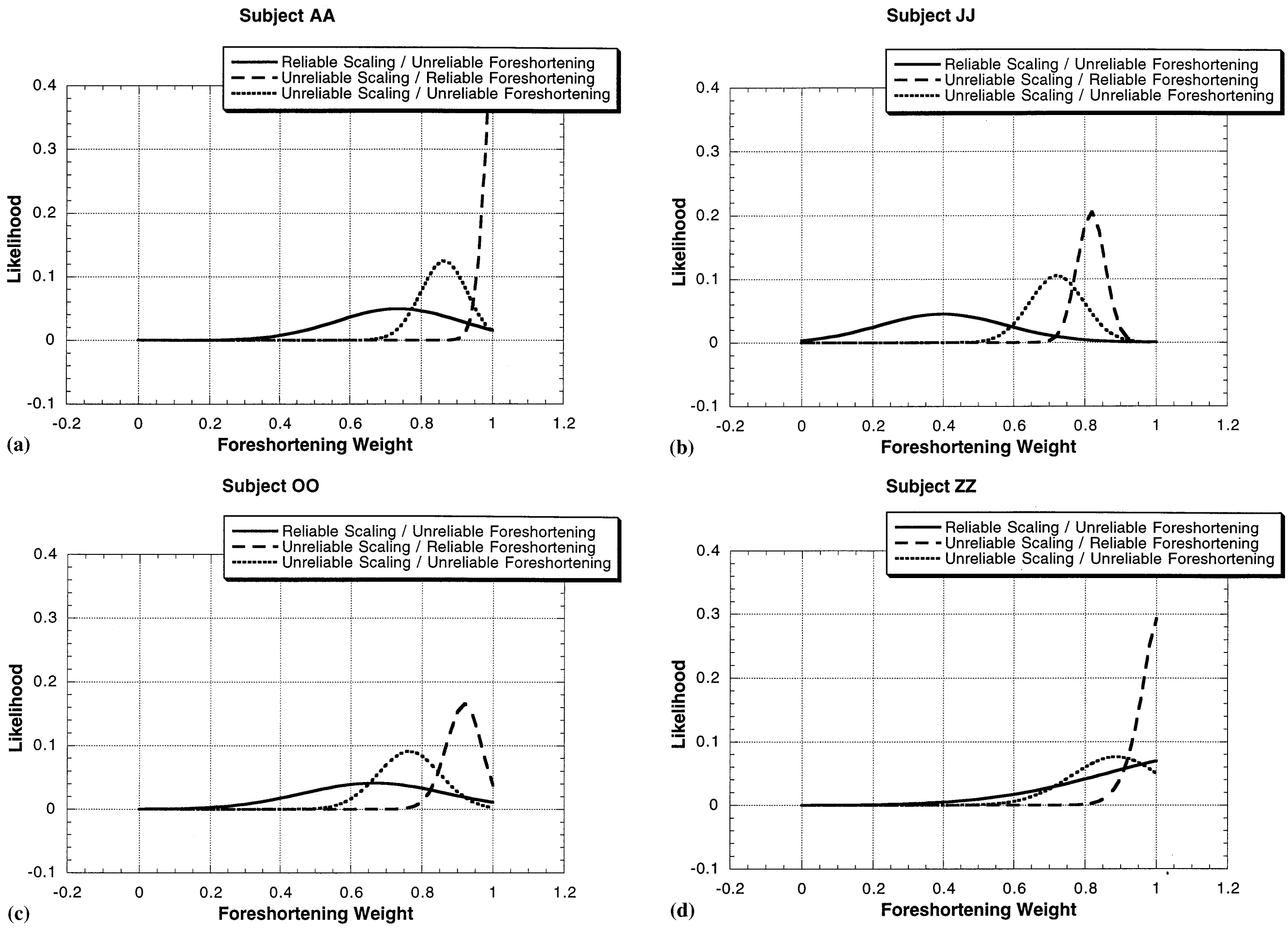


Fig. 6. Likelihood functions for cue weights plotted for each of the four subjects in experiment 1.

4.3. Discussion

4.3.1. Estimation bias

Before discussing the theoretical import of the results, we must consider a technical issue regarding the estimation procedure whether or not it is biased in any way. In no condition were the perturbation levels in the scaling and foreshortening cues commensurate. Moreover, the perceptual interpretation of the different cues is corrupted by early level sensory noise of unknown magnitude, possibly further exacerbating the effective difference in perturbation levels. These differences can lead to biases in some weight estimation methods [18]; therefore, it was important to ascertain that our results were not contaminated by such biases.

We ran a number of simulations of model systems in which the data was collected from 'experiments' with an observer that performed linear weighted averages of cue slants to make discriminations. We manipulated independently the variance of cue perturbations and the simulated levels of the observers' weights for different experiments. In none of the simulated cases did the maximum likelihood approach to fitting observers' weights to discrimination data give rise to any estimation bias. That is to say, the expected value of the estimated weights was equal to the true observer weights regardless of the relative perturbation levels of the cues.

4.3.2. Texture cue weights

The most striking result of the analysis is that subjects clearly relied heavily on foreshortening information to perform the slant discrimination task. There is some hint that the relative weighting of foreshortening information decreases when scaling information is most reliable; however, even here the weight assigned to foreshortening is significant. The significant role of foreshortening in the discrimination task is consistent with previously cited evidence that this cue is of primary importance in the perception of planar surface slant from stochastic textures [3]. Conflicting evidence generally comes from experiments employing regular texture patterns [2,3,5] suggesting that cue weights may be qualitatively different for highly regular texture patterns. As Frisby et al. have argued, the large weights given to scaling information for regular textures may result from the use of a linear perspective cue which does not appear in images of stochastic textures.

Table 1
Maximum likelihood estimates of foreshortening weight, with position weight set to zero

Subject	Condition	Foreshortening weight	95% Confidence interval
AA	Rel. Scaling/Unrel. Foreshortening	1	(0.45, 1)
AA	Unrel. scaling/Rel. Foreshortening	1	(0.88, 1)
AA	Unrel. scaling/Unrel. foreshortening	0.89	(0.66, 1)
JJ	Rel. scaling/Unrel. foreshortening	0.74	(0.39,0.99)
JJ	Unrel. scaling/Rel. foreshortening	1	(0.95, 1)
JJ	Unrel. scaling/Unrel. foreshortening	0.87	(0.72,0.99)
OO	Rel. scaling/Unrel. foreshortening	0.67	(0.25,0.97)
OO	Unrel. scaling/Rel. foreshortening	0.92	(0.82, 1)
OO	Unrel. scaling/Unrel. Foreshortening	0.76	(0.58,0.92)
ZZ	Rel. Scaling/Unrel. Foreshortening	0.4	(0.06,0.77)
ZZ	Unrel. Scaling/Rel. Foreshortening	0.81	(0.73,0.9)
ZZ	Unrel. Scaling/Unrel. Foreshortening	0.72	(0.56,0.86)

(MLE estimates of position weights were all zero, with three exceptions very close to zero- subject AA had position weights of 0.05, 0.04, and 0.05 none of which were significantly greater than 0. The scale weight is given by one minus the foreshortening weight.

The heavy reliance of subjects on texture foreshortening information to make judgements of surface slant raises the question of how they interpret the cue. In particular, we must ask whether they rely on a prior assumption that surface textures are isotropic to interpret texture foreshortening.

5. Experiment 2: Testing for isotropy

Using an isotropy assumption to interpret foreshortening information would have a number of obvious advantages. It would make the estimation of surface orientation from texture much more local than could be supported by a more general assumption of homogeneity (which requires using spatial 'gradients' in texture foreshortening). It also renders texture foreshortening information significantly more reliable, when the isotropy assumption is correct. For images of isotropic texture patterns like those used in experiment 1, ideal observer thresholds for discriminating surface slant increase by a factor of three to five times when prior knowledge of isotropy is removed from the ideal observer [12].

In this section, we describe an experiment and associated ideal observer perturbation analysis designed to test whether or not subjects rely on an assumption of surface texture isotropy to discriminate surface slant from texture. In the experiment, we used a variety of types of surface textures to test how generalizable the role of isotropy is. These include elliptical element textures, Voronoi textures and rectangular element textures. The rectangular element textures contain a strong 'higher-order' shape cue to surface orientation which is not conditioned on a putative prior assumption of isotropy the skew of projected texture elements. We included these textures in the experiment to test for an influence of higher-than-second order texture shape information on human perception of slant from texture.

5.1. Strategy

As in the previous experiment, we measured subjects' ability to discriminate surface orientation from texture. In one condition of the experiment, we used stimuli generated from isotropic surface textures. In other conditions, we randomly perturbed the 'degree' of anisotropy of the surface textures used to generate stimuli. To do this, we globally compressed isotropic surface textures by random amounts, in random directions, before projection into a stimulus image. The resulting stimulus sets simulated truly random samples drawn from an ensemble of surface textures with varying amounts of global compression. These included isotropic textures, but only as a special case. Fig. 7 illustrates the method.

The logic of the experiment follows from the observation that anisotropies in surface texture patterns will induce predictable biases in the perceived orientation of a surface for a system which assumes surface texture isotropy. Fig. 8 illustrates the biasing effects of globally compressing surface textures before projection; an observer which uses only foreshortening information and assumes isotropy will overestimate surface slant for images of surface textures compressed in the direction of surface tilt and will underestimate slant for images of surface textures compressed in the perpendicular direction. Because of these biasing effects, the random variations in surface texture compression which we have used to make the anisotropic stimuli for the experiment create large perturbations in the surface slant indicated by what we will refer to as the foreshortening-withisotropy cue. The foreshortening-with-isotropy cue is simply the foreshortening cue interpreted using an assumption of surface texture isotropy.

We can use the foreshortening ideal observer derived using the isotropy assumption to measure the perturbations in the foreshortening-with-isotropy cue. Correlating subjects' discrimination judgements on a trial-by-trial basis with the stimulus slants suggested by the foreshortening-with isotropy ideal observer provides a means to measure the strength of subjects' assumptions of isotropy.

We followed a two-stage strategy in our analysis. First, we performed a non-parametric test to determine if the proportion of times subjects' judged the comparison stimulus to have a greater slant than the test stimulus was reliably correlated with the slant difference indicated by the foreshortening-with-isotropy cue. Second, we applied a variant of the linear model described above to estimate weights assigned to different texture cues, including the foreshortening-with-isotropy cue.

The linear model used to predict subjects' data included, as before, scaling, position and foreshortening-with-isotropy cues. We further included in the analysis a 'null' cue, which was set to the true slant difference for each trial; thus, the model contained four weight parameters (constrained to sum to one). The null cue served as a dummy variable to represent the degree to which subjects' perceptual estimates of slant were pulled away from the biased interpretations derived from the isotropic ideal observer and not accounted for by the scaling and position cues. One factor which would contribute to a significant weight for the null cue would be a less than complete assumption of isotropy; that is, a soft rather than hard bias towards isotropy. We will consider other factors which could contribute to this weight in the context of the results obtained from the analysis.

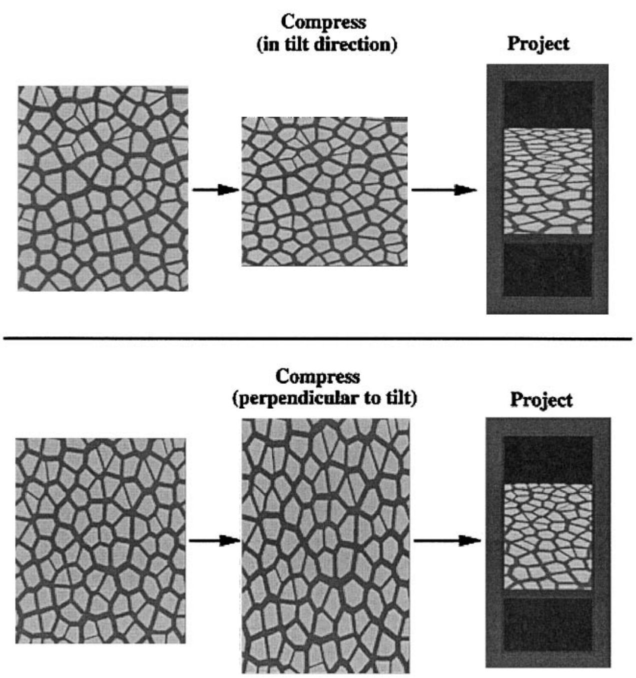


Fig. 7. Stimuli for the experiment were created in three steps. First a random, isotropic texture was created (Examples of Voronoi textures are shown here). Second, the texture was compressed by a random factor between 0.7 and 1.0. In the figure, the compression was done either in the direction of surface tilt (top figure) or in the perpendicular direction (bottom figure). In the experiment, the direction of compression was chosen randomly between 0 and 180°. Finally, the texture was mapped onto a planar surface viewed at a fixed slant away from the line of sight. In the figure, both textures are mapped onto surfaces slanted away from the line of sight by 60°. If you have a bias to interpret the surface textures as isotropic, then you should see the top figure as being more slanted away from you than the bottom figure.

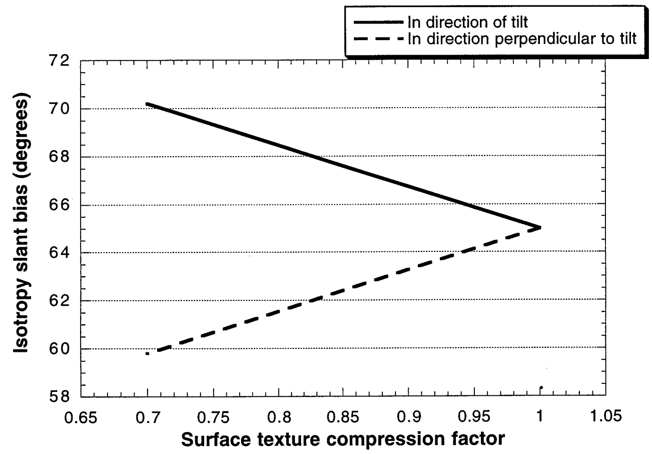


Fig. 8. The average slant estimate derived from the foreshortening ideal observer which assumes isotropy, as a function of the factor by which a surface texture is globally compressed away from being isotropic (in the direction of surface tilt). Compression factors less than one represent greater compression, compression factors greater than one represent stretching of the texture. Small compression factors represent more compression. The true slant is 65°. The range of compression factors is the same as used in the experiment; thus, the range of biases predicted by isotropy varies from $+5$ to $-5°$. Stimuli in the experiment were created by compressing surface textures not only in the tilt direction, but also in randomly chosen directions between 0 and 180°. The predicted slant biases for the experimental stimuli, assuming the tilt is known, depends both on the compression factor and the direction of compression. in the experiments, the horizontal boundaries of the stimulus surfaces provided a strong cue, as confirmed by subject reports, that the surfaces were tilted in a vertical direction.

5.2. Methods

5.2.1. Apparatus

We used the same apparatus for experiment 2 as we used for experiment 1. We refer the reader to the methods section for experiment 1 for details.

5.2.2. Stimuli

Four classes of textures were used to create stimuli for the experiment: isotropic, elliptical element textures; anisotropic, elliptical element textures; anisotropic, Voronoi textures (see figure lb); and anisotropic, rectangular element textures. The anisotropic textures were generated by globally compressing surface textures by random factors between 0.7 and 1.0 before projection. The compression was applied in randomly selected directions over the range 0–180°. We chose to do this, rather than always compressing textures in the tilt/anti-tilt directions, in order to maximize the likelihood of subjects detecting the anisotropy of the stimulus ensemble and switching to an anisotropic interpretation strategy (the tilt direction is reliably determined by the horizontal orientation of the surface boundaries in the image). Thus, our estimates of the strength of the isotropy assumption will not be overly optimistic, as might occur were we only to compress textures in directions which more directly mimic the effects of projective foreshortening in the known tilt direction.

A special method was used to create anisotropic rectangular element textures. We did so by first compressing an elliptical element texture and then replacing the ellipses with rectangles of equivalent length, orientation and aspect ratio. This gave rise to textures whose global orientation statistics were the same as the elliptical element textures, but whose elements were composed of rectangles.

The elliptical element, Voronoi and rectangular element textures were equated to have the same statistics in their second-order spatial moments. We accomplished this by first generating Voronoi textures from constrained random lattices of sample points, and then computing the statistics of the ellipses fitted to the Voronoi polygons in the texture patterns (see [11,12] for details). The Voronoi texture statistics were used to create the elliptical element and rectangular element textures. This implies that if one were to replace the texels in the Voronoi and rectangular textures with their fitted ellipses, they would appear qualitatively the same as the elliptical element textures. The one exception to this is that the Voronoi textures show no overlap due to the constrained way in which they were created.

The remainder of the stimulus parameters (size, number of texture elements and average size of elements) were the same as in the previous experiment.

5.2.3. Procedure

The procedure used for the experiment was the same as that used in the previous experiment.

5.2.4. Subjects

Subjects were drawn from the student body at the University of Pennsylvania and were paid for their participation. Subjects had normal or corrected to normal vision and were naive to vision science.

5.3. Results

5.3.1. Discrimination thresholds

Fig. 9 shows 75% thresholds for the four conditions of the experiment, computed in the same way as in experiment 1. The comparisons of interest are between the isotropic elliptical and anisotropic elliptical conditions and between the anisotropic elliptical and the two other anisotropic texture types. For three of the four subjects, discrimination thresholds are significantly higher for the anisotropic elliptical textures than for the isotropic elliptical textures. No consistent pattern emerges across the three anisotropic texture conditions.

The hypothesis that subjects employ some form of isotropy constraint to infer surface orientation from texture predicts that discrimination performance should decrease when the degree of anisotropy in surface textures is randomly perturbed in the experiment. This is broadly consistent with the significant difference in thresholds found for three subjects between the isotropic and anisotropic elliptical texture conditions. The result does not, however, prove the point. Other factors may well cause stimuli generated from anisotropic textures to be less discriminable. We therefore turned to a closer, trial-by-trial comparison of subjects' judgements with the stimulus slant differences suggested by the foreshortening-with-isotropy cue.

5.3.2. A qualitative test of isotropy

In our first pass at rigorously testing whether subjects relied to some degree on an assumption of isotropy to make judgements of surface slant in the experiment, we wanted to avoid making any assumptions about the form of subjects' psychometric functions or of the cue-integration process (e.g. linearity). We therefore performed a non-parametric test on the data obtained in the three anisotropic texture conditions. The null hypothesis was that within each fixed level of comparison-test slant difference no significant correlation would be found between the proportion of times subjects judged the comparison stimulus to be most slanted and the slant difference indicated by the foreshortening-with-isotropy cue (as measured by application of the isotropic, foreshortening ideal observer). The isotropy hypothesis, on the other hand, predicts that the greater the slant difference indicated by foreshortening-with-isotropy, the greater the proportion of times subjects would judge the comparison stimulus to be most slanted.

In order to test the null hypothesis, we first binned trials, at each level of comparison-test slant difference, into three bins. These corresponded to smaller, medium and larger differences in slant as measured by the isotropic, foreshortening ideal. The boundaries of the bins at each level of comparison-test slant difference were determined so as to equate, as much as possible,

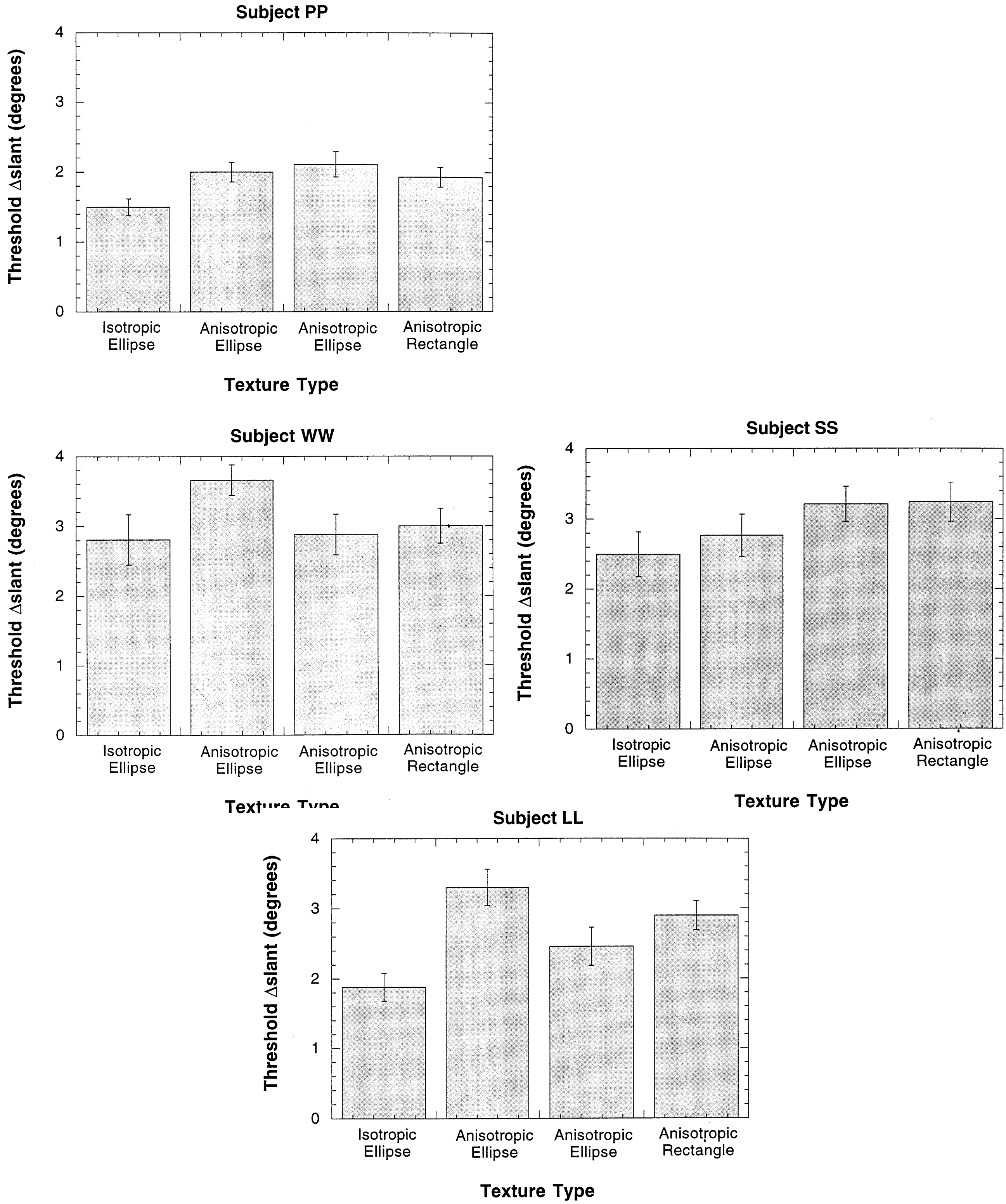


Fig. 9. Threshold data from the second experiment. For simplicity, we have only plotted one-sided discrimination thresholds; that is the positive difference in slant between test and comparison stimuli needed to correctly judge the comparison stimulus to have greater slant 75% of the time.

the number of trials falling into each bin. We then performed a χ^2 test for the goodness-of-fit of the null model, which predicts flat psychometric functions within each level of comparison-test slant difference.

Fig. 10 shows the resulting, binned data for subject PP in the three anisotropic texture conditions. Plots for the other three subjects show a similar pattern. The effect's of an apparent assumption of isotropy are (see Table 2). Out of the twelve comparisons (four subjects, three anisotropic texture conditions), only three fail to reach a significance level of 0.05, though two of these come close. The analysis confirms that in most texture conditions, subjects' discrimination performance correlates with the slant difference indicated by the foreshortening with-isotropy cue, for fixed physical slant differences.

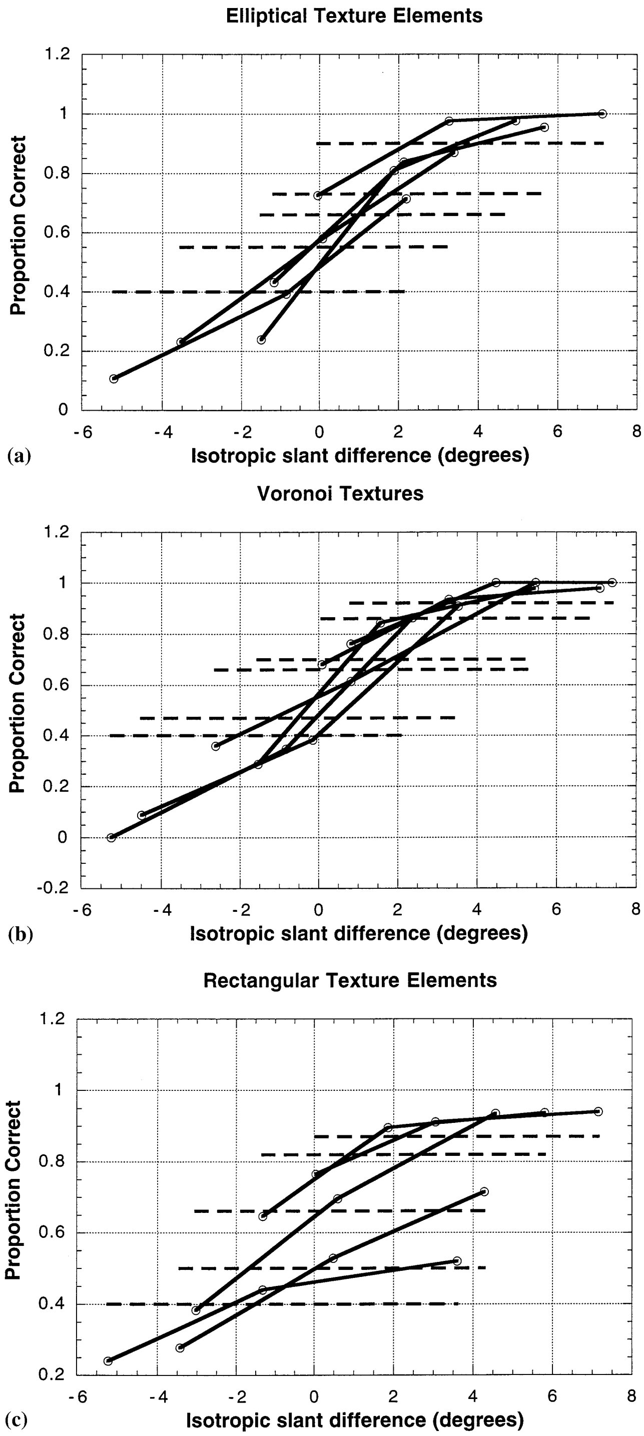


Fig. 10. Plots of the percentage of times subject PP selected the comparison stimulus as having greater slant than the test stimulus as a function of the slant difference indicated by the foreshortening-with-isotropy cue (after binning). The dashed lines indicate the results predicted by the null hypothesis. For each level of slant difference, these lines are flat, at the level of the average percentage of trials in which subjects choose the comparison stimulus to have greater slant (labeled here as percent correct). The width of the lines reflects the range of perturbations in the foreshortening-with-isotropy cue induced by the random perturbations in surface texture compression.

The approach just described has a number of problems. First, by its non-parametric nature, it provides only a qualitative test of isotropy. We would ideally like some measure of the strength of subjects' isotropy bias. Second, simulations of the ideal observers for the scaling and position cues revealed that compressing surface textures before projecting them into an image induced biases in these cues which were correlated with the biases induced in the foreshortening-with-isotropy cue. This is not surprising for the position cue, which gains some of its informativeness from the relative spacing between texels, and can therefore by biased by an isotropy assumption in much the same way as the foreshortening cue[2]. It is more surprising for the scaling cue, which we would expect to be free of assumptions about the orientation statistics of a texture. In fact, the dependence of the scaling cue on surface texture compression depends on what image measurements serve as the basis of the cue. In our case, we have equated scaling information with the information provided by the spatial distribution of texel lengths in an image, which, it turns out, leads to small biasing effects of surface texture compression[3]. While the scaling bias is weak (it is approximately half that of the foreshortening bias for the range of compressions used here, and decreases to zero for larger compressions), it could theoretically explain the results of the analysis.

The difficulties in interpreting the non-parametric test led us to perform a more quantitative analysis in which we computed the weights given by subjects to position, scaling and foreshortening with-isotropy cues in their discrimination strategies.

[2] We used the position ideal observer derived with an assumption of isotropy for the simulation.

[3] Compressing a surface texture composed of texels at random orientations by small amounts has the effect of increasing the variance of texel lengths. The scaling ideal observer effectively selects estimates of slant which minimize the variance of back-projected surface texel lengths. Since the scaling ideal observer knows nothing about the surface texture compression, it gains some advantage in picking slants for which the surface texture compression is 'undone' by the decompressing effects of back-projection. The bias disappears for compression factors much less than 0.5, but this is smaller than the range of compressions used in the current experiment.

Table 2
χ^2 values for testing whether the probability of judging the comparison stimulus to have greater slant than the test stimulus is unaffected by an isotropy assumption

	Texture Type		
Subject	Anisotropic Elliptical	Anisotropic Voronoi	Anisotropic Rectangular
PP	$x^2(18) = 66.7(<0.0005)$	$x^2(18) = 91.9(<0.0005)$	$x^2(16) = 29.65(<0.01)$
WW	$x^2(18) = 34.42(<0.01)$	$x^2(18) = 34.06(<0.01)$	$x^2(18) = 9.08(*)$
SS	$x^2(18) = 44\ 59(<0.0005)$	$x^2(20) = 25.13(*)$	$x^2(20) = 26.89(*)$
LL	$x^2(20) = 54.45(<0.0005)$	$x^2(16) = 60.08(<0.0005)$	$x^2(16) = 30.2(<0.01)$

The numbers in parentheses are the degrees of freedom for the test. Only levels of slant differences which were tested more than fifteen times in the experiment (giving a minimum of five trials per bin) were used for the test. A * indicates that the level of significance was above 0.05.

5.3.3. Linear perturbation analysis

We used the technique described in Section 3.3 to estimate the weights given by subjects to different texture cues in the current experiment. The list of cues we included in the estimation procedure included scaling information, foreshortening (with isotropy) information, position (with isotropy) information and the null cue described above, which is simply the true comparison-test slant difference. The reason for inclusion of the null cue will become clear shortly.

The foreshortening cue defined here is sub-ideal in the sense that it relies on the incorrect assumption of isotropy. We would like the model to accommodate the possibility that the visual system shows a graded bias towards isotropy. One way to do this would be to simulate foreshortening observers which simultaneously estimate global surface texture compression along with surface slant. A graded bias towards isotropy would be modeled in this framework as a prior distribution which assigns graded probabilities to different surface texture compression parameters. The difficulty in this technique is largely practical searching for the best fitting parameters for such a prior distribution is computationally prohibitive. We therefore chose to include within the linear formulation a dummy cue which represents perfect knowledge of stimulus slant. As described in the introduction to this section, a stronger prior assumption of isotropy should lead to less weight being given to the null cue in the analysis.

Other facts could contribute to a stronger weight being given to the null cue. The most interesting of these, for the Voronoi and rectangular textures, is the presence of other information in the stimuli not accounted for by the scaling, foreshortening and position ideal observers, which only make use of the second-order spatial moments of texels (length, orientation and aspect ratio). For Voronoi textures, some information is provided by the perspective convergence of parallel texel borders (probably a minor effect given the limited size of the texels), and other information exists in the higher-order spatial moments of the texels. For rectangular element textures, texel skew provides a reliable cue to surface slant independent of texel size, orientation and aspect ratio. The influence of these alternative sources of information should be reflected in the weight given to the null cue.

As in the first experiment, we found no cases in which significant weight was given by subjects to the position cue. For ease of exposition, therefore, we have removed this cue from the model fits. Table 3 summarizes the maximum likelihood estimates of the remaining three cue weights obtained from the analysis. To show the results graphically, we plotted the estimates of weights computed for each subject in a triangular space representing the possible relative combinations of weights, constrained to some to 1 (see Fig. 11). The vertices of the triangles correspond to one or another of the cues being given a weight of 1. Points along the left side of the triangle correspond to a zero scaling weight, points on the right side correspond to a zero foreshortening-with-isotropy weight and points along the bottom correspond to a zero null weight. The center of mass of the triangular space is the point at which all three weights are equal.

The data clearly show only a small weight being given to the scaling cue across all subjects and texture types. The foreshortening-with-isotropy cue receives a strong weight for all subjects and conditions, but is reduced significantly relative to the null cue for rectangular textures.

5.4. Discussion

5.4.1. Isotropy

The data not only confirms the results of the first experiment, that subjects gave minimal weight to the scaling cue to make discriminations; it also clearly reveals a strong isotropy assumption in subjects' interpretations of slant from texture. The strength of the assumption is similar for both the elliptical element

Table 3
Maximum likelihood estimates of the weights given to scaling, foreshortening-with-isotropy and the null cue (see text for discussion)

Subject	Condition	Scaling weight	Foreshortening-with-isotropy weight	Null weight
PP	Ellipse	0.23	0.62	0.15
PP	Voronoi	0.09	0.68	0.23
PP	Rectangle	0.2	0.35	0.45
WW	Ellipse	0	0.78	0.22
WW	Voronoi	0.06	0.57	0.37
WW	Rectangle	0	0.41	0.59
SS	Ellipse	0.03	0.68	0.29
SS	Voronoi	0.09	0.53	0.38
SS	Rectangle	0	0.54	0.46
LL	Ellipse	0.11	0.89	0
LL	Voronoi	0.15	0.82	0.03
LL	Rectangle	0.09	0.48	0.43

The three stimulus conditions shown are the anisotropic stimulus conditions, in which the degree of anisotropy of surface textures (amount by which they were globally compressed) was randomly perturbed before projection into stimulus images.

textures and the more natural Voronoi textures. It is reduced somewhat for the rectangular element textures.

The most straightforward way to interpret the meaning of the linear model presented here is to use the combination of cue weights to predict biases in perceived slant as a function of surface texture compression. We will refer to these biases as isotropy biases. In particular, for a given image, we have for a prediction of perceived surface slant the equation, where σ is the perceived slant for a stimulus, σ_s is the slant indicated by the scaling information in the stimulus, σ_f is the slant indicated by the foreshortening-with-isotropy information and σ_n is the actual slant of the stimulus[4]

$$\sigma = w_s\sigma_s + w_f\sigma_f + w_n\sigma_n \tag{15}$$

The bias induced by compressing a surface texture before projection into the image is given by

$$\Delta\sigma = w_s\Delta\sigma_s + w_f\Delta\sigma_f \tag{16}$$

where $\Delta\sigma_s$ and $\Delta\sigma_f$ are the biases induced by the scaling and foreshortening-with-isotropy cues in the stimulus, respectively. Since we know from simulations that the scaling cue bias is approximately half that of the foreshortening-with-isotropy bias for anisotropic textures, we can express the overall bias as a factor of the bias predicted by the foreshortening-with-isotropy cue. Writing the predicted bias as $\Delta\sigma = \alpha\Delta\sigma_f$, we obtain a factor of $\alpha = 0.5 + w_f$. Applying this equation to the weights derived from the current experiment, we find predicted bias factors ranging from 0.41 to 0.95. This is consistent with the wide range of biases found in other experiments [2,8].

[4] We should properly apply the linear model in the variance-stabilized slant domain for which we derived estimates of weights. We have found, however, that the resulting predicted biases are well-approximated by applying the simple linear model in the standard slant domain.

We should emphasize that the current experiment mixed a random selection of surface texture compression factors and directions of compression. The estimated weights are therefore some form of average of the weights for the different surface texture compressions. It may well be that the isotropy bias is more pronounced in some conditions than others; for example, conditions in which the compression is in the direction of tilt. We will consider this point further in the general discussion.

5.4.2. Higher-order texture shape information

The predicted bias factors for the anisotropic elliptical and rectangular texture types are listed in Table 4. The data for all four subjects shows that the isotropy bias is significantly weaker for rectangular element textures than it is for the elliptical element textures. Since the texture stimuli in these conditions had identical information in the second-order spatial moments of the textures, the result is consistent with the hypothesis that the visual system uses higher-order texel shape information such as that provided by local texture skew. This is true even though the textures used here were stochastic not regular rectangular grids for which one might naturally expect skew to have a significant impact on perceived slant.

6. General discussion

6.1. Foreshortening-with-isotropy

The results of the analysis are consistent with the hypothesis that subjects rely most heavily on local foreshortening information (i.e. with the isotropy assumption) to judge surface slant from texture, even for images of extended planar surfaces. This contradicts the

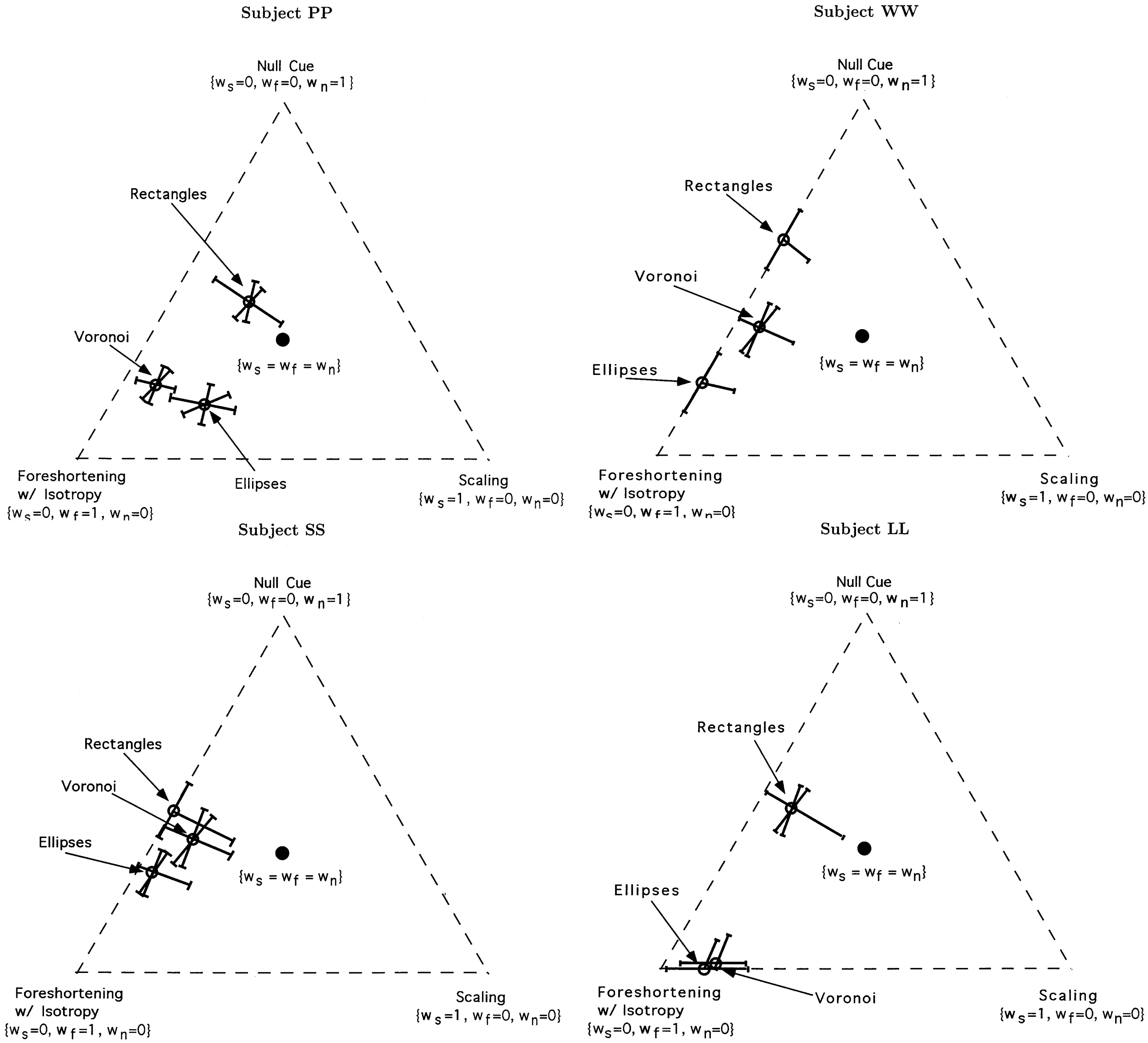


Fig. 11. Maximum likelihood estimates of cue weights calculated for each of the four subjects in experiment 2. The sides of the triangle correspond to one or another weight (corresponding to the facing vertex) being given zero weight. Position along a side, or any parallel line interior to the triangle, specifies the proportion of weight given to the cue associated with the vertex at one end of the side relative to the cue associated with the vertex at the other end. Error bars were estimated from the second derivative of the likelihood function computed in the direction of each vertex, and were truncated at the boundaries of the space.

conclusions of [6], but is consistent with a growing body of literature on the perception of planar surface orientation from texture [2,3,8]. The finding of a strong reliance on foreshortening-with-isotropy is also consistent with results on human perception of surface curvature from texture, both in small field of view displays [4,9,13,14] and in large field of view displays [3,6]. The conclusion, however, is tempered by the finding of a significant difference between predicted isotropy biases and the actual biases estimated from the experimental data in experiment 2. While it is not entirely clear to what to attribute this difference, it does indicate that subjects integrated a local isotropy strategy with a more global gradient-based strategy, since, at least in the ellipse condition, it is only gradient-based information which provided unbiased information about surface orientation.

The textures composed of randomly oriented rectangles elicited less of an isotropy bias than did the textures composed of ellipses or, for three of the subjects, the Voronoi textures. This is consistent with the hypothesis that higher-order local texture moments (skew) contribute to subjects' orientation judgements. More generally, it highlights the dependence of observers' estimation strategies on texture type. Another example of this is the finding that for textures composed of regular arrays of rectangles (arrayed in parallel), subjects give more weight to scaling information than to foreshortening [3]. Frisby et al. have suggested that the contradiction between this result and other results suggesting the dominance of foreshortening information may be explained by the fact that the regular arrays of rectangles contained a strong linear perspective cue, a global cue which does not appear in images of randomly oriented texture elements.

Decomposing texture information into cues contained in the local second-order spatial moments of texture patterns and those contained in higher-order moments provides a route to reconciling the results obtained with different texture types. The existing data is best fit by the hypothesis that second-order spatial moment information is interpreted by giving most weight to foreshortening information with a strong assumption of isotropy, but that higher-order spatial cues (linear perspective, skew) contribute significantly to the percept of surface orientation when present in images. Whether such higher-order shape cues contribute to curvature perception remains to be seen.

6.2. Other texture cue decompositions

We have decomposed texture information in a very particular way: into scaling, foreshortening and position cues. While this decomposition has a sound computational basis (e.g. separating the scaling and foreshortening effects of projection), one must consider the possibility that the visual system relies on a different decomposition. This is particularly relevant to the cues which rely on spatial gradients in texture properties; in particular, scaling and foreshortening-without-isotropy. Todd et al. [13], for example, have suggested that the visual system uses a form of foreshortening-with-isotropy to measure local surface tilt, but uses the derivative of texel widths (computed in the tilt direction) to measure local surface curvature in the direction of tilt. This idea, per se, is problematic, since the spatial gradient of texture widths is dependent on both curvature and slant [19][5], but it does make clear the possibility of other parameterizations of texture cues.

It is true that the weight given to the scaling cue may in fact reflect a reliance on a texture measurement other than texel length (e.g. texel width computed perpendicular to surface tilt, or texel area); however, the principle result of the current study is that subjects relied most heavily on the foreshortening-with-isotropy cue. This cue is essentially local and cannot be approximated by combinations of other forms of texture gradients. Thus, the results do directly implicate the foreshortening-with-isotropy cue as we have defined it.

Table 4
Estimates of the slant bias factors derived from the weights estimated in each of the ellipse and rectangle texture conditions of the experiment

	Isotropy bias factor		
Subject	Ellipse Textures	Rectangle Textures	Z-test
PP	0.735 ± 0.065	0.45 ± 0.063	$Z = 3.15$ ($p < 0.001$)
WW	0.78 ± 0.086	0.41 ± 0.084	$Z = 3.12$ ($p < 0.001$)
SS	0.695 ± 0.063	0.54 ± 0.07	$Z = 1.65$ ($p < 0.05$)
LL	0.955 ± 0.085.	0.525 ± 0.071	$Z = 3.88$ ($p < 0.0001$)

The bias factor specifies fraction of the foreshortening-with-isotropy bias which should appear in subjects' estimates of slant for the stimuli used in the experiment. The confidence intervals on the bias factors are the estimated standard errors of the estimates, derived from the Hessians of the likelihood functions (effectively fitting Gaussian distributions to the likelihood functions in the neighborhood of the peak). These estimates were used to compute the Z-scores for the difference in bias factors between texture conditions.

6.3. When to use isotropy and when not to

The results of the second experiment implicate a soft form of the isotropy assumption. While this can be implemented in several ways, there is an 'ideal' way to soften the isotropy assumption. In particular, the assumption should be embodied in a mixture of experts model [20] in which the likelihood function for surface orientation from texture is derived from a mixed prior model of surface textures. This is a particular type of mixture-of-experts model which Yuille and Clark have refered to as a competitive prior model [21]. We briefly discuss here the structure and implications of such a model. We will assume for simplicity that surfaces are constrained to be planar. In practice, such a constraint may derive from other cues (e.g. straight surface boundaries) present in an image.

A reasonable assumption about natural surface textures is that, while not all textures are isotropic, some non-zero percentage are. This leads to a mixed prior model of surface textures, with some percentage being isotropic and some not. The likelihood function for such a model is a weighted sum of two likelihood functions,

$$p_f(\vec{T}/\sigma,\tau) = \pi\, p_{\text{isotropic}}(\vec{T}/\sigma,\tau) + (1-\pi)p_{\text{anisotropic}}(\vec{T}/\sigma,\tau) \tag{17}$$

where $\vec{T}$ represents a collection of measurements of the local shape and orientation properties of an optical texture pattern, σ and τ represent slant and tilt, respec-

[5] The psychophysical evidence cited in support of the hypothesis may also be interpreted in other ways, as discussed in Cumming, et. al.[14]

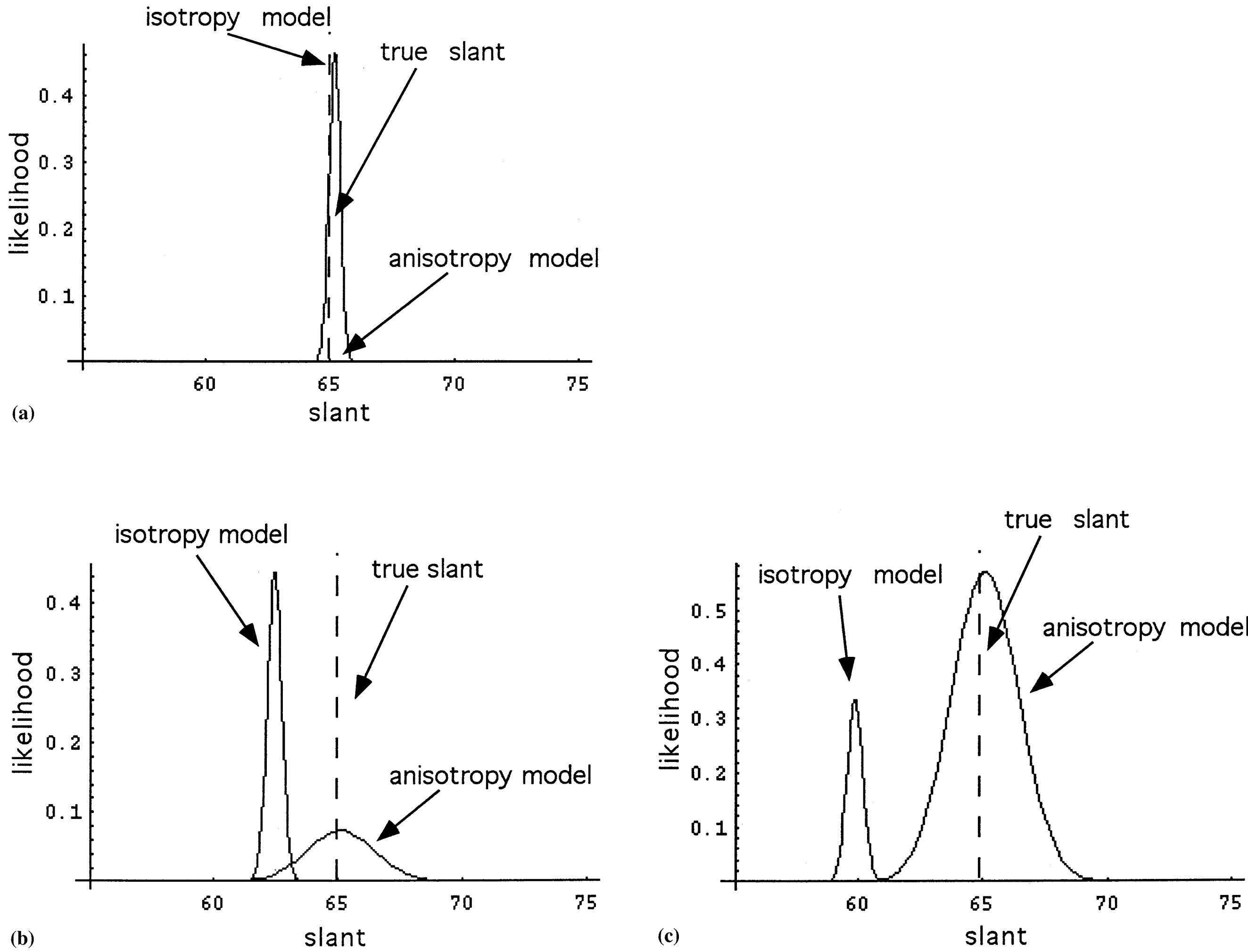


Fig. 12. Likelihood functions for the isotropic foreshortening and anisotropic foreshortening cues. The former was derived assuming that surface textures are isotropic. The latter was derived assuming that all surface texture compression factors and directions of compression were equally likely; that is, that the global 'aspect ratios' of surfaces textures was uniformly distributed from between O and 1 and the global orientations of the textures were uniformly distributed between O and 2π. (a) For isotropic textures, the magnitude of the likelihood function for the anisotropy model is too small relative to the isotropy likelihood function to appear on the graph; (b) For a surface texture compression of 0.85 in a direction perpendicular to the surface tilt, the isotropy model still dominates; (c) For a compression factor of 0.7, the anisotropy model becomes weakly dominant. Note the increasing bias in the isotropy model with increasing surface texture anisotropy as well as the decrease in the goodness of fit of the isotropy model.

tively and π is the proportion of surface textures in the world assumed to be isotropic (the likelihood functions are implicitly conditioned on surface planarity, which may itself have some attached probability). The relative magnitudes of the two component likelihood functions determine whether or not to impose isotropy, or how much to weight the isotropy interpretation of surface orientation.

The central insight to seeing how this might work is the observation that the likelihood function for the isotropic sub-model will decrease in magnitude for images of anisotropic textures, since the global texture pattern will not be consistent with a planar, isotropic surface texture. This allows the information contained within the foreshortening cue itself to automatically determine whether or not the isotropy assumption is appropriate for a given stimulus, when other information is available to constrain a surface to be planar. To illustrate how this might work, we have derived example likelihood functions for the isotropic model and an anisotropic model (in which the orientation and magnitude of global surface texture compression are both assumed to be uniformly distributed) for stimuli like those used in experiment 2. Fig. 12 shows sample results for the image of an isotropic surface texture and images of the same surface texture compressed by factors of 15% and 30% in a direction perpendicular to surface tilt.

Two features of the mixed likelihood function are notable. First, for truly isotropic textures, the isotropy sub-model clearly dominates the mixed likelihood function; thus, the foreshortening information contained in sample images of isotropic textures is enough by itself to determine that isotropy is an appropriate assumption to use for texture interpretation. Second, the relative heights of the two component likelihood functions change with increasing anisotropy of the surface texture used to create an image greater anisotropy leads to a shrinking of the isotropic likelihood function relative to the anisotropic likelihood function. When the compression factor is much greater than 30%, the magnitude of the isotropy likelihood function becomes vanishingly small relative to the anisotropic likelihood function (the actual transition point will depend on how regular textures are, field of view and so on).

The mixed likelihood function specifies in an objective way the information content of the foreshortening cue in a texture pattern. How to select a slant based on the foreshortening information is an open question. When the magnitudes of the component likelihood functions differ greatly, any reasonable strategy will reduce to switching to one or another strategy. When the magnitudes of the likelihood functions are more similar, a variety of strategies might be used selection of the interpretation corresponding to the model with highest peak likelihood, random selection between the modes, with appropriately weighted probabilities (as some have argued is done for other bimodal stimuli such as Necker cubes) or a weighted average of the modal interpretations. All of these strategies would reflect themselves in subjects' showing a decreasing isotropy bias as a function of the amount of surface texture compression.

In the present experiment, the range of texture compression factors used was quite small, the largest being 30%. For such stimuli, the experimental finding of a strong, but less than total isotropy bias is entirely consistent with a rational strategy for interpreting foreshortening information. The analysis of the mixture model does suggest some interesting predictions, however. It predicts that the isotropy bias should degrade with increasing compression of surface textures. It also predicts that field of view size should have a significant effect on the degree to which isotropy biases interpretations. The larger the field of view, the greater the information in the texture pattern that isotropy does not apply. Finally, the analysis suggests a strongly non-linear way in which other cues may interact with texture information. In particular, other cues such as stereo or motion could selectively support one or another of the isotropy and anisotropy interpretations. This effect would appear in the joint likelihood function for texture and other cues, given by the product of the individual likelihood functions for texture and the other cues. Were the modes of the likelihood functions for other cues to be near one or another of the foreshortening modes, it would selectively enhance that mode.

Our discussion of using foreshortening without assuming isotropy is predicated on the assumption that subjects see surfaces as being planar, and in our experiments, subjects generally reported seeing the surfaces as planar. Without this assumption, or some similar assumption of surface smoothness, the foreshortening cue by itself would not provide enough information to adequately constrain surface shape interpretation without assuming isotropy. This suggests that scaling information may contribute more heavily to shape from texture perception when the visual system 'turns off' isotropy. The most general computational model of shape from texture currently available, developed by Malik and Rosenholtz, does not assume isotropy or any other constraint on the spatial distribution of textures other than homogeneity. The model implicitly uses all the available cues in a texture pattern to make inferences about surface shape [22]. Even using all the available cues, they find that they must impose some form of local smoothness constraint (in their case, an assumption that a surface's gaussian curvature locally constant), to constrain the reconstruction problem.

Whether the human visual system can interpret shape from texture when textures are clearly anisotropic is questionable [14], except in cases in which higher-order texture information like skew or linear perspective is available. The latter cues, of course, rely on constraints on the spatial structure of surface textures similar to isotropy; skew relies on a bilateral symmetry constraint and linear perspective relies on a parallelism constraint. This observation suggests that further work needs to be done to understand all of the constraints underlying human perception of surface shape and orientation from texture. A complete model of human perception of shape from texture will likely require a more complex mixture of experts model which includes all the prior constraints which human observers may impose to interpret shape from texture.

6.4. Why not use scaling information?

Scaling information was given only a small weight in the stimulus conditions used in the experiments. This weight might well increase under large field viewing conditions; however, it is not a-priori clear that it should—the relative reliability of foreshortening and scaling information does not change that much with increasing field of view size (in the range from 12.5 to 30°) [12]. This result applies to the theoretical limits on cue reliability. In practice, the dependence of scaling cue reliability on field of view size will depend as much on the visual system's ability to resolve spatial differences in size as it does on the information content in the stimulus.

One computational argument for using scaling information is that many textures are not 'insurface', as were the ones used in these experiments. Scaling information retains the same qualitative form for other types of textures (e.g. rocks, grass), while the nature of foreshortening information changes dramatically. Elongated objects perpendicular to a surface, for example, exhibit decreasing foreshortening with increasing surface slant, rather than the increasing foreshortening exhibited by in-surface textures. Scaling information may take on a greater perceptual weight for stimuli composed of such textures than is found for in-surface textures.

6.5. Future directions

The ideal observer methodology described here provides a powerful analytical tool for measuring cue weights and for testing specific models of cue interpretation. Our results have gone a long way towards quantifying how humans judge planar surface slant from texture, but many questions remain. We must further measure performance for large field of view surfaces. This will require using texture patterns projected onto real surfaces, as carried out by [2,3,7], to minimize conflicts with other cues such as accommodation and blur. Furthermore, the micro-structure of how the visual system imposes the isotropy constraint remains to be clearly elucidated. The question of how the visual system determines when to apply an isotropy constraint is central to this problem. More generally, the structure of the system of constraints imposed by the visual system on surface textures, including constraints such as bilateral symmetry and parallelism, needs to be elucidated.

7. Summary

We have introduced a novel technique for measuring cue weights and for comparing different functional models of cue interpretation. The technique is a form of perturbation analysis in which the natural perturbations in cue content are taken advantage of by back-correlating subjects judgements on a slant discrimination task with slant estimates derived from different ideal observers. These sub-ideals use only one or another of the available cues or rely on specific prior assumptions not necessarily valid for all of a set of experimental stimuli. Perhaps the greatest power of the method is its ability to test different functional models of cue interpretation, as in the test of whether subjects imposed an isotropy constraint to perceive surface slant in the experimental stimuli.

We found that subjects relied primarily on the foreshortening cue to discriminate slant from texture and that they relied to varying degrees on an assumption of isotropy. We also obtained evidence that the visual system can use more than just second-order spatial moments of texture patterns to perceive slant, even for stochastic textures without any global structure. In particular, subjects appeared to be able to rely in part on the skew of rectangular texture elements to make judgements of surface slant for textures composed of randomly shaped and oriented rectangles.

Acknowledgements

I am grateful to Ginny Richards for many helpful discussions on perturbation analysis and to Paul Schrater, Jack Nachmias and Mary Bravo for helpful discussions concerning this work. This work was supported by NIH grant NEI-EY09383 to the author.

References

[1] Gibson JJ. The perception of visual surfaces. Am J Psychol 1950;63:367–84.
[2] Frisby JP, Buckley D, Freeman J. Experiments on stereo and texture cue combination in human vision using quasi-natural viewing. In: Orban HH, editor. Artificial and Biological Visual Systems. Berlin: Springer-Verlag, 1992.
[3] Frisby JP, Buckley D, Freeman J, Stereo and texture cue integration in the perception of planar and curved large real surfaces, in:Inui T, McClelland JL,(Eds.), Attention and Performance XVI: Information and Integration in Perception and Communication. Cambridge: MIT Press, 1996.
[4] Young M, Landy M, Maloney L. A perturbation analysis of depth perception from combinations of texture and motion cues. Vis Res 1993;33(18):2685–96.
[5] Braunstein ML, Payne JW. Perspective and form ratio as determinants of relative slant judgments. J Exp Psychol 1969;81(3):584–90.
[6] Cutting J, Millard R. Three gradients and the perception of flat and curved surfaces. J Exp Psychol Gen 1984;113(2):198–216.
[7] Buckley D, Frisby J, Blake A. Does the human visual system implement an ideal observer theory of slant from texture? Vis Res 1996;36(8):1163–76.
[8] Rosenholtz R, Malik J. Shape from texture: Isotropy or homogeneity (or both)? Vis Res 1997;37(16):2283–94.
[9] Blake A, Marinos C. Shape from texture: estimation, isotropy and moments. Artif Intell 1990;45:323–80.
[10] Blake A, Bulthoff HH, Sheinberg A. Shape from texture: Ideal observers and human psychophysics. Vis Res 1993;33(12):1723–37.
[11] Knill DC, Discriminating surface slant from texture: Comparing human and ideal observers, Vis Res.(in press)
[12] Knill D, Surface orientation from texture: Ideal observers, generic observers and the information content of texture cues, Vis Res.(in press)
[13] Todd J, Akerstrom R. Perception of three-dimensional form from patterns of optical texture. J Exp Psychol 1987;13:242–55.
[14] Cumming BG, Johnston EB, Parker AJ. Effects of different texture cues on curved surfaces viewed stereoscopically. Vis Res 1993;33(5–6):827–38.
[15] Landy MS, Maloney LT, Johnston EB, Young M. Measurement and modeling of depth cue combination: in defense of weak fusion. Vis Res 1995;35(3):389–412.

[16] Kendall MK, Stuart A. Design and analysis, and time series. The Advanced Theory of Statistics, vol. 3. New York: Hafner, 1966.
[17] Press WH, Vetterling WT. Numerical Recipes in C: The Art of Scientific Computing. New York: Cambridge University Press, 1992.
[18] Richards VM, Zhu S. Relative estimates of combination weights, decision criteria, and internal noise based on correlation coefficients. J Acoust Soc Am 1994;95(1):423–34.
[19] Garding J. Shape from texture for smooth curved surfaces in perspective projection. J Math Imaging Vis 1992;2(4):327–50.
[20] Jordan MI, Jacobs RA. Hierarchical mixtures of experts and the EM algorithm. Neural Computation 1994;6(2):181–214.
[21] Yuille AL, Clark JJ. Bayesian models, deformable templates and competitive priors. In: Harris L, Jenkin M, editors. Spatial Vision in Humans and Robots. England: Cambridge University Press, 1994.
[22] Malik J, Rosenholtz R. Recovering surface curvature and orientation from texture distortion: A least squares algorithm and sensitivity analysis. Proceedings 3rd European Conferance on Computer Vision, Volume 800 of Lecture Notes in Computer Science. Berlin: Springer-Verlag, 1995.