PERGAMON

Vision Research 38 (1998) 1683–1711

Vision Research

Discrimination of planar surface slant from texture: human and ideal observers compared

David C. Knill *

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

Received 23 October 1996; received in revised form 31 July 1997

Abstract

In order to quantify the ability of the human visual system to use texture information to perceive planar surface orientation, I measured subjects' ability to discriminate planar surface slant (angle away from the fronto-parallel) for a variety of different types of textures and in a number of different viewing conditions. I measured the subjects' discrimination performance as a function of surface slant, field of view size and surface texture structure. I compared the subjects' performance with that of ideal observers derived for each of the available texture cues—texel position, scaling and foreshortening. The results can be summarized by four points: (i) subjects' discrimination performance improves dramatically with increasing surface slant, tracking the performance of the ideal observers; (ii) subjects can integrate texture information over a large range of visual angles; (iii) comparisons between human subjects and ideal observers show that the human observers rely to some degree on foreshortening information; and (iv) similar comparisons show that in using foreshortening information, subjects rely to some extent on a prior assumption of isotropy.

Keywords: Surface slant; Texture; Ideal observers; Foreshortening; Isotropy

1. Introduction

1.1. General

Texture patterns provide useful cues to the 3D spatial layout of surfaces in our environment. Fig. 1, for example, shows texture patterns which induce strong percepts of planar surfaces receding in depth. Psychophysical research on the perception of planar surface orientation from texture has focused on three questions; how strong of a 3D cue is texture [1–5], which component cues within texture patterns (scaling, foreshortening and density) does the visual system use to estimate surface orientation and curvature [6–14], and what prior assumptions about surface textures does the visual system rely on to interpret texture cues; particularly, does the visual system assume isotropy [10,15]? What has not been studied are the basic limits on the ability of the human visual system to use texture information to make judgments about surface shape and orientation. In this paper we describe a series of experiments designed to systematically explore these limits.

1.2. Qualitative 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 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

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

0042-6989/98/$19.00
PII: S0042-6989(97)00325-8

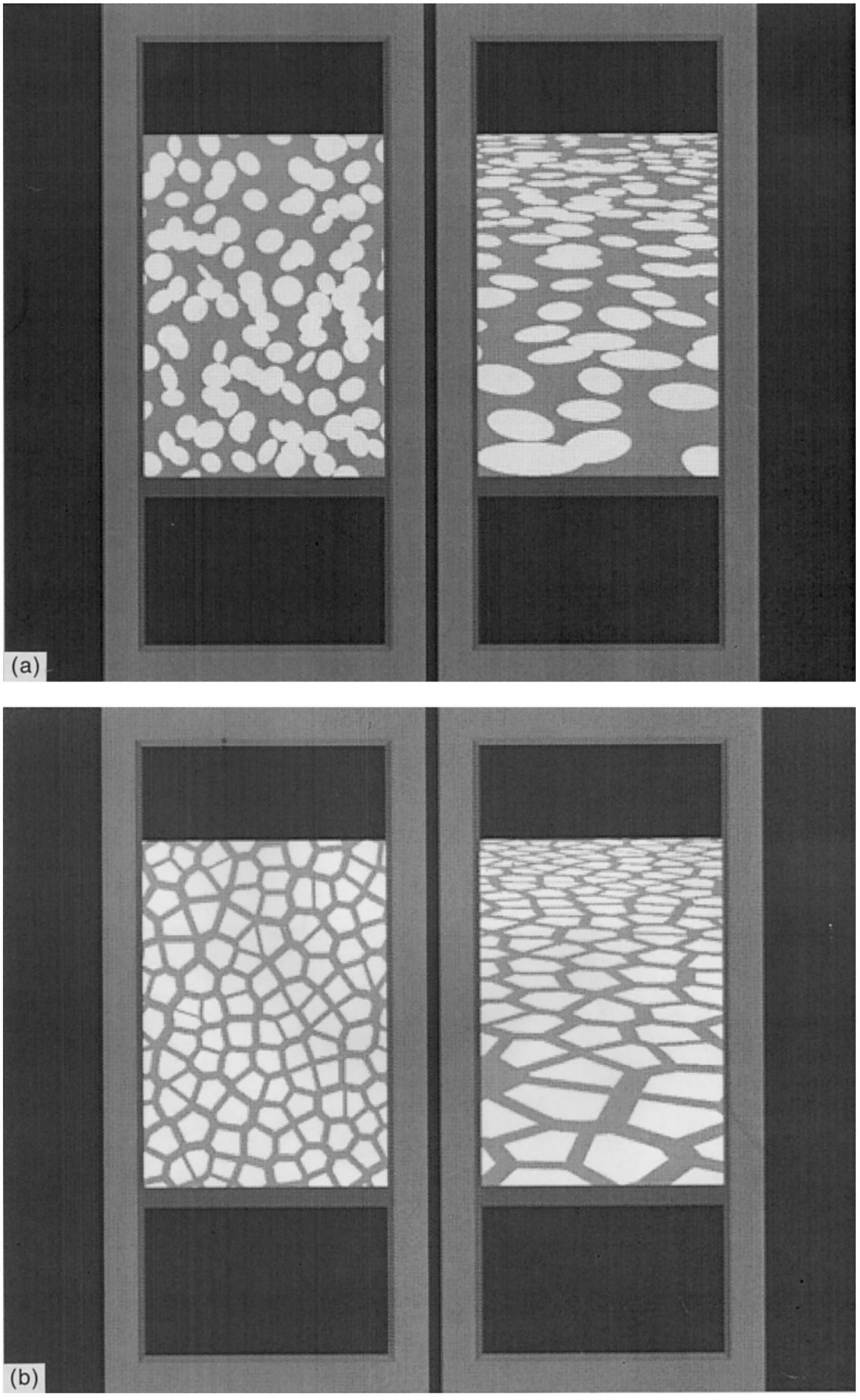


Fig. 1. Examples of texture patterns which induce strong percepts of surfaces slanted in depth. In each of the two figures, one surface is fronto-parallel and the other is slanted away from the line of sight at 73°. Images like these were used as stimuli for the experiments.

available to an observer, who only sees the effects of these distortions in the projected image. A-priori 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 judgments about surface orientation. With it, an observer could make orientation judgments 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 [16]), 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 [8,9,11,12]. Some authors', however, have defined the compression cue differently, as the gradient in absolute compression of texture elements in the direction of surface tilt [10,14]. Compression defined this way depends on both scaling and foreshortening effects of projection (and does not admit a straightforward application of the isotropy constraint). We chose to use the term, foreshortening, to avoid confusion and to emphasize the nature of the cue—the distortion in texture element shapes induced purely by projective foreshortening.

1.3. Goals and organization of paper

In the experiments described here, we measure the subjects' abilities to discriminate planar surface slant (angle away from the fronto-parallel) in a variety of stimulus conditions. The work is motivated in part by the results of ideal observer analyses of texture information, which have helped to elucidate the computational structure of texture information for surface slant. Following the example of Blake et al. [12], in their study of curvature perception from texture, we will compare human discrimination performance to that of 'provisional' ideal observers, each of which makes optimal use of one of the three texture cues to make discrimination judgments. The provisional ideal observers, taken together, completely characterize the information content of the stimuli used in the experiments described here. These comparisons will provide some insight into how the visual system uses texture information to make inferences about planar surface slant.

The second section of the paper reviews the results of previously reported ideal observer analyses. Section 3 described the general methodology used for all of the experiments described here. The remaining sections describe the results of five experiments. In experiment 1, we look at how the subjects' discrimination performance varies with surface slant. In experiments 2 and 3 we measure the effects of field of view on discrimination performance. In experiment 4, we independently vary the reliability of the information provided by the scaling and foreshortening cues with an eye towards measuring the relative contributions of the different cues to the subjects' slant judgments. Experiment 5 tests the generalizeability of the results to more naturalistic textures than those used in the first four experiments.

2. Ideal observer results

Several authors have derived ideal observers for shape-from-texture which are provably optimal estimators of surface geometry from texture (for specific classes of texture) [12,16–20]. The ideal observers are defined as maximum likelihood estimators of surface geometry parameters (e.g. planar surface orientation) from a given set of image texture measurements. The most recent and general of the formulations uses as texture information the positions and second-order spatial moments of discrete element textures [16]. This formulation of the ideal observer is optimal for textures composed of elliptical texture elements embedded in planar surfaces (Fig. 1a)[1]. We will, therefore, use elliptical element textures as stimuli for our experiments.

We derived four provisional ideal observers for estimating surface orientation from texture, which, combined, make use of all of the information available in elliptical element textures [16]. The set of provisional ideal observers includes one each for the scaling and position cues, and two for different variants of the foreshortening cue, one which relies on a prior assumption of surface texture isotropy and one which does not. We will occasionally refer to the former as the 'isotropic' foreshortening ideal observer and the latter as the anisotropic foreshortening ideal observer. We performed a wide range of ideal observer simulations to explore the properties of texture information as it applies to the inference of planar surface orientation [16]. The general results of these and previous ideal observer simulations applied to estimating planar surface orientation [12] may be summarized as follows:

- Texture information becomes more reliable as surface slant increases.
- The relative reliability of the three texture cues depends on the size of one's field of view on a surface and on the statistics of surface textures. For large fields of view (e.g. $>20°$), in which all three cues are reasonably reliable, the relative reliability of the cues depends primarily on surface texture statistics. One can manipulate the relative reliability of the cues by varying these statistics—increasing the variance of surface texel sizes decreases the reliability of the scaling cue, etc. Some insight into the relative reliability of the cues in nature may be induced from studying stereotypical textures. Simulations of both maximum entropy textures (the most 'random' possible textures) and a naturalistic class of random tile textures (Voronoi textures [15], as shown in Fig. 1b) show a common pattern in which foreshortening information is consistently more reliable (at slants $>30°$) than scaling information, and both are more reliable than position information.
- Texture information derived from further points on a surface (i.e. the tops of ground-plane images) contributes disproportionately to the overall information content of stimuli. This interacts with the general increase in informativeness which arises from having a broader field of view to determine the effects of the field of view on the texture cue informativeness. In general, when the field of view is decreased from the bottom-up (for ground plane images), texture cue reliability decreases much more slowly than when the field of view is decreased from the top-down.
- Even with large fields of view, the reliability of foreshortening information depends heavily on the isotropy assumption. The foreshortening ideal observer derived without a-priori knowledge of surface texture isotropy performed significantly worse than the ideal observer which had knowledge of isotropy, when applied to images of isotropic textures.

[1] The ideal observer derived by Blake et al. [12] is optimal for surface textures composed of uniformly and independently positioned and oriented line elements.

3. General methods

We begin by describing aspects of the experimental methodology which were common to all experiments. We also briefly describe the method used to estimate ideal observer performance in the same experimental conditions as subjects.

3.1. Apparatus and viewing set-up

Stimuli were presented on the display monitor of an SGI computer. The monitor was an SGI model TFS6705, 17 inch, 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. This eliminated stereo cues to flatness and reduced as much as possible the motion-parallax cues, while keeping the viewing situation reasonably comfortable for subjects. The subjects' non-viewing eye was covered with an eye-patch to eliminate any potential for binocular rivalry. The subjects were tested in a room painted matte black to minimize secondary reflections back onto the monitor (another flatness cue). Finally, a matte black occluder was placed over the front of the monitor to obscure the physical screen boundaries.

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° × 40° of visual angle. The actual visual angle of stimuli depended on experimental conditions. 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.

3.2. Stimuli

Fig. 1 shows examples of the stimuli used in the experiments. The stimuli were computer-rendered images of surfaces with textures composed of randomly positioned, sized and shaped texels. Stimulus images for each trial were generated in three stages: surface texture generation, projection into the image and cropping of the texture image to generate an image of a spatially bounded, slanted surface viewed through a window frame.

- Surface texture generation: For each experiment, a particular setting of surface texture model parameters was selected to define a stochastic ensemble of textures to use for generating stimuli. Stimuli were randomly sampled from this ensemble on each trial. Prior to running the experiments, a set of 1000 samples of a 2-D stochastic diffusion process were used to generate constrained random lattices of texel positions (Appendix A). Each sample consisted of an 800 × 800 element random lattice of points in a unit square. On a given trial, the first step in generating the surface texture was to randomly select a lattice of texel positions and scale and translate it so that the prescribed number of texels (± 2%) appeared in the displayed surface area. A texel was considered to be 'in' the display area if its center of mass appeared within the area. This approximately fixed the number of texels shown in the stimuli for a particular experimental condition. At each of the resulting surface positions, texels were 'drawn' on the surface by randomly selecting texel sizes, orientations and aspect ratios from the prescribed texture model.
- Projection: After generating a surface texture, the texture was projected into an image according to the perspective projection model described in ref. [16]. The model maps surface texel parameters to image texel parameters. After projection of the texels into an image, the sizes of the texels were scaled to fix the total display area covered by texels on a given surface (generally at 50%), eliminating total texel area or any of its derivatives (e.g. image contrast) as a cue for the discrimination task. After computing the projected image texel parameters, texels were drawn as 20-sided polygonal approximations to ellipses. At this resolution, no effects of the discretization were apparent in the display. Texel luminance was fixed at 150 gray-scale units and the background was fixed at 60 gray-scale units, giving reasonably high contrast texture patterns.
- Cropping: In the projection phase of stimulus generation, texels appearing outside the desired surface boundaries were projected into the image. The resulting surface image was cropped on the sides by the vertical edges of a simulated window frame and at the top and bottom by the horizontal edges of the surface. This resulted in small portions of texels appearing at the boundaries of the surface displays which were not included in the count used to fix the number of texels within the display area. The addition of these texel fragments to the predetermined number of texels fit into the display area is likely to have been offset by the loss of information from those other texels where they intersected the boundary of the display area. At the bottom edge of the displayed surfaces, we drew a 30-pixel wide lip, giving the surface the appearance of having thickness. Besides disambiguating the direction of surface slant (not a problem in most of the experimental conditions), the lip added a phenomenally significant degree of realism to the displays.

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.

3.3. Procedure

We used a two-alternative forced choice procedure in which subjects judged which of the 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 judgments 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 in some of the experiments (except experiment 5) 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.

Two psychophysical methods were used to estimate thresholds and psychometric functions; the method of constant stimuli and a non-parametric stair-case method (used when we were measuring both discrimination thresholds and points of subjective equality between different classes of texture stimuli). In both cases, we made maximum likelihood fits of psychometric function parameters to the raw data to derive estimates of discrimination thresholds. Thus, the non-parametric staircase served primarily as a sampling procedure in the final analysis.

Before starting the main part of the experiment, the 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 the test and target stimuli, respectively). Besides serving to show the subjects the task they were to perform, which they picked up immediately, this phase of an experiment served to weed out the subjects who did not see any 3D effect from the texture stimuli. Out of more than 50 subjects we have run in these and other similar experiments, only two such subjects were found. Within each experiment, experimental conditions were randomly interleaved in each session, and sessions were designed to last for no more than 1 h each. Depending on the experiment, the subjects ran for anywhere between 8 and 16 sessions.

3.4. Data analysis

Pilot data showed that the subjects' thresholds decreased markedly with increasing surface slant (by a factor of 10 or more over the range of test slants used). This suggested that the variance of perceptual estimates of surface slant decrease monotonically with increasing surface slant. This can lead to significant skewing of the subjects' psychometric functions. In order to model this asymmetry, we chose a non-standard technique for modeling the psychometric function. We assumed that a one-parameter transformation of surface slant could be found which stabilized the variance of the subjects' perceptual estimates and which normalized the probability distribution of the estimates. This allowed us to apply a standard Gaussian signal discrimination model to subjects' data in the transformed slant domain, with the psychometric function in this domain characterized cumulative Gaussian. The form of the variance stabilizing function was derived by assuming than the standard deviation of the subjects' slant estimates was linear in the neighborhood of the test slant. This led to a log transform [21] of slant with the generic form (Eq. (1))

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

where σ_0 is the test slant. We will refer to the parameter, β, as the stabilizing transform. The full psychometric model can then be expressed as a two-parameter function in the slant domain of the form,

$$p(\text{correct}) = erf(\log_e[1 + \beta(\sigma - \sigma_0)]/s) \tag{2}$$

where s is the slope parameter for the error function. In our analysis, we characterized psychometric functions by their 85% thresholds, T, and their variance stabilizing parameters, β. These we estimated using standard maximum likelihood parameter fits to the function given in Eq. (2)

Standard deviations for psychometric parameter estimates were derived using Fisher's asymptotically correct approximation for the parameters' covariance matrix [22]. This is given by (Eq. (3))

$$\Sigma^{-1} = -\begin{pmatrix} \dfrac{\partial^2 L}{\partial T^2} & \dfrac{\partial^2 L}{\partial T \partial \beta} \\ \dfrac{\partial^2 L}{\partial T \partial \beta} & \dfrac{\partial^2 L}{\partial \beta^2} \end{pmatrix} \tag{3}$$

where L is the likelihood function for T and β conditioned on the subjects' data. The standard deviation of the subjects' thresholds in this case would be given by $\sigma_T = \sqrt{\Sigma_{1,1}}$

3.5. Ideal observer simulations

In [16] the standard deviation of ideal observers' estimates of surface slant as a summary measure of texture information about surface orientation. To make the ideal observer data more commensurate with our human data, we ran simulated experiments with the ideal observer under the same conditions as used for the human experiments. In the ideal observer experiments, we used an adaptive psychophysical procedure (Quest [23]) based on a Weibull function model of the psychometric function to estimate the ideal observers' 85% threshold for discriminating surface slant from texture[2]. Stimuli were constructed for the ideal observer

[2] We have modified the Quest procedure to work with other psychometric functions, but have found that the choice of psychometric function has no significant effect on the resulting estimates of ideal observer thresholds.

simulations in exactly the same way as for the human experiments. For each experimental condition, ten experimental sessions of 150 trials each were run to estimate discrimination thresholds. The threshold slant differences to which the Quest procedure converged at the end of each session were averaged to obtain the final threshold estimates.

4. Experiment 1

The first experiment measured the subjects' thresholds for discriminating surface slant from large field-of-view textures over a range of test slants.

4.1. Methods

4.1.1. Stimuli

Stimuli consisted of images of textured surfaces subtending a width of 500 pixels and a height of 624 pixels. The images of surfaces were framed as described in the general methods section above. At the viewing distance of 28 cm which we used, this gave images of surfaces whose textured regions subtended $20° \times 25°$ of visual angle. Surface textures were generated from the following model.

4.1.1.1. Texel lengths. The lengths of surface texels were independently and randomly sampled from a Gaussian distribution with a standard deviation of 0.2 and a mean of 1. They were then scaled by an amount necessary to insure that texels covered 50% of a stimulus image. This simulated a texel length process with a floating mean (determined implicitly by the viewing conditions) and a standard deviation which was 20% of the mean.

4.1.1.2. Texel orientations. As described above, surface texel orientations were independently and randomly sampled from a uniform distribution over the range, $(0, 2\pi)$.

4.1.1.3. Texel aspect ratios. The aspect ratios of texels were independently and randomly sampled from a distribution qualitatively similar to truncated Gaussian distribution with a mean parameter of 1 and a standard deviation parameter of 0.25. The actual distribution was defined by Eq. (4)

$$p(a_i) = k \exp\left[-\left(\log\left(\frac{a_i}{1-a_i}\right) - 0.8\right)^2 / 2\right]\left(\frac{1}{a_i - a_i^2}\right) \quad (4)$$

Justification for this particular form of distribution is given by Knill [16].

4.1.1.4. Texel positions. Texel positions were created as described in the general methods section, that is, a set of random position lattices was generated from a stochastic-diffusion process and randomly sampled to create the set of texel positions for each trial (Appendix A).

Fig. 2 shows examples of the stimulus pairs used in the experiment.

4.1.2. Procedure

We used a method of constant stimuli to estimate psychometric functions for surface slant discrimination for five different test slants, 0, 30, 50, 65 and 70° (measured from the center of the display to the observation point). The resulting data was used to derive estimates of one-sided discrimination thresholds. For each test slant, five target slant differences (all positive) were estimated from pilot studies to evenly sample the psychometric function at that slant. The subjects ran in 16 sessions each. Each session consisted of 15 trials per condition randomly interleaved through the session, giving a total of 375 trials per session. The first session was discarded from the data analysis as practice.

4.1.3. Subjects

Four undergraduates naive to the purposes of the experiment and naive to vision science in general served as subjects for the experiment. All had normal or corrected to normal vision.

4.2. Results

Fig. 3 shows plots of the estimated 85% thresholds for each of the four subjects along with plots of the thresholds for each of the four provisional ideal observers simulated. For the particular texture parameters used to generate the stimuli for this experiment, foreshortening information (assuming prior knowledge of isotropy) is the most reliable cue to surface slant, followed by scaling and position information. The foreshortening ideal observer which does not use prior knowledge of the stimulus textures' isotropy performs significantly worse than one which has such knowledge reflecting the significant role played by an assumption of isotropy in making the foreshortening informative.

In the general methods section, we described the steps taken to insure that subjects did not rely on simple 2D strategies for performing the discrimination task (removing simple first-order cues, limiting feedback to occasional percent correct scores, etc.). These precautions, while necessary, could not insure that subjects did not use simple 2D discrimination strategies; therefore, we interviewed the subjects after they finished the experiment to gauge the possible influence of 2D strategies. All the subjects in this and the other experiments reported that the stimuli elicited phenomenally

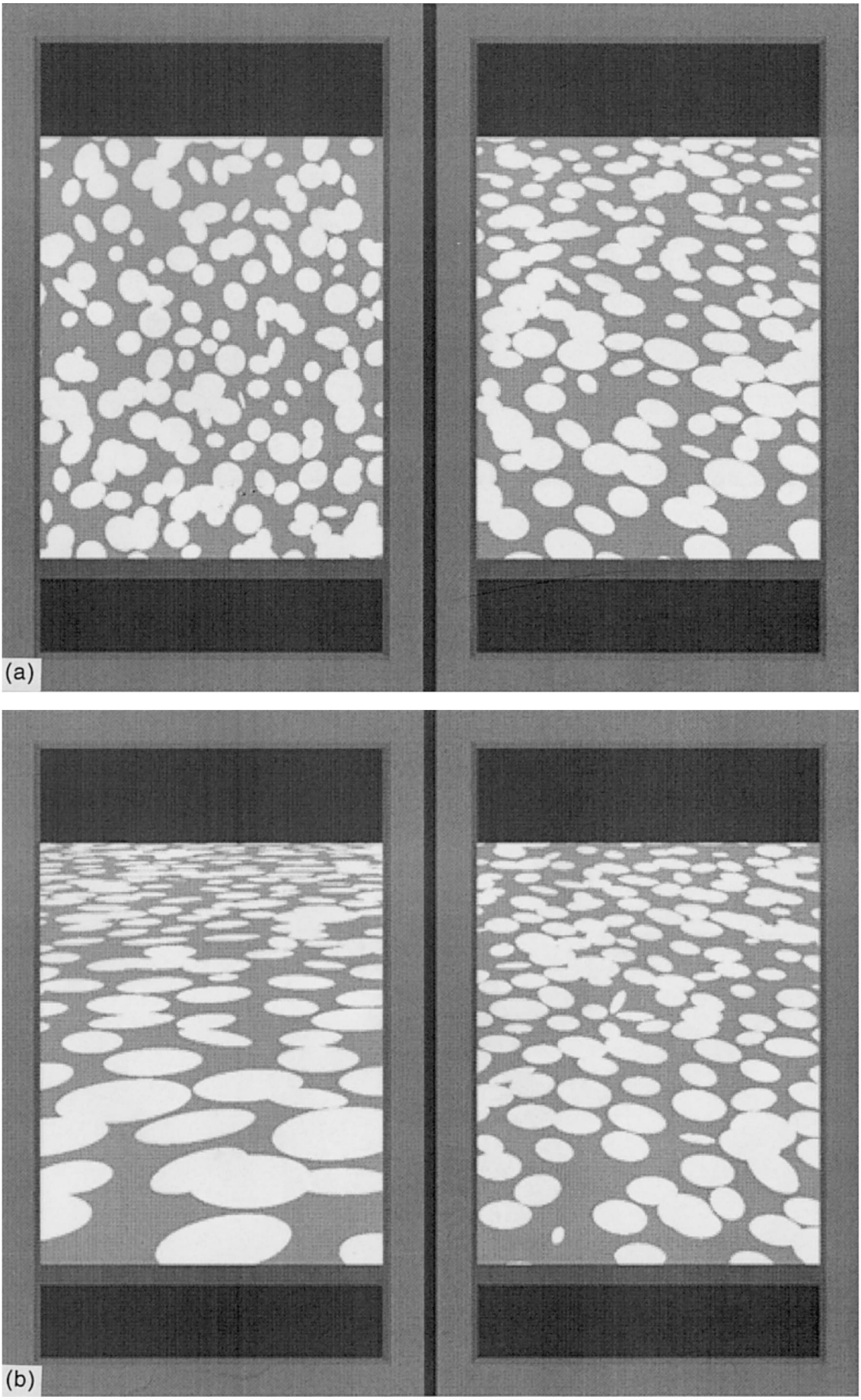


Fig. 2.

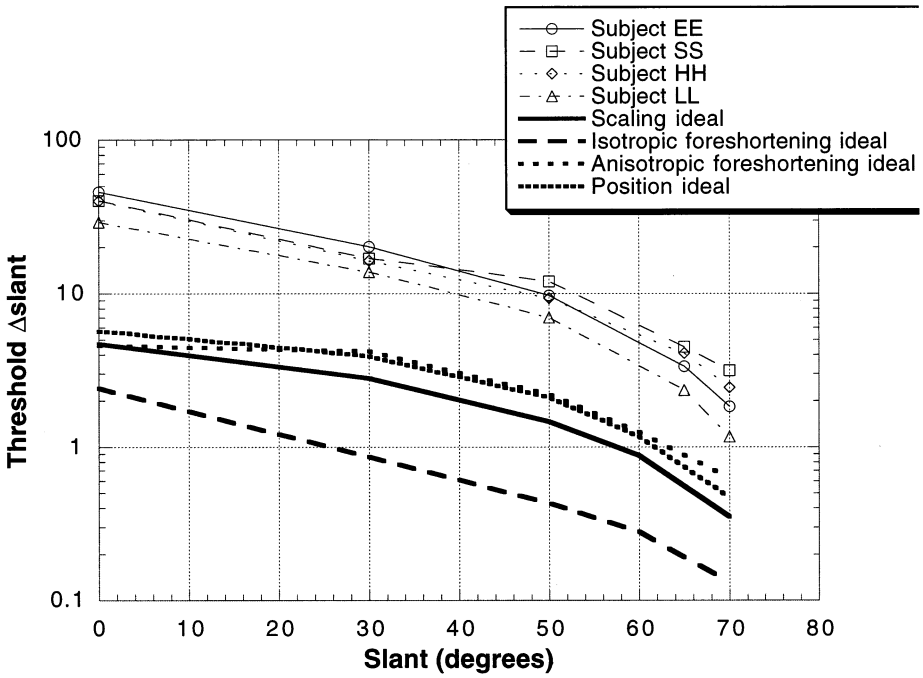


Fig. 3. A log-linear plot of the subjects' thresholds and of the thresholds for the four provisional ideal observers described in the text. Error bars derived from the likelihood function for threshold were less than the size of the markers used in the plot. The isotropic foreshortening ideal observer is an ideal observer which has prior knowledge that the surface texture ensemble used to create stimuli is isotropic. The anisotropic ideal observer does not make such an assumption.

strong percepts of planar surfaces receding in depth. When asked about what strategies they used to perform the experiment, subjects uniformly reported either explicit 3D strategies (e.g. selecting the surface which would be the easiest to walk up or the one whose top was further away relative to the bottom) or no apparent strategy other than to select which stimulus appeared phenomenally more slanted.

4.3. Discussion

4.3.1. Dependence of threshold on absolute slant

The plots of ideal observer thresholds reflect the changing reliability of the three texture cues as a function of surface slant. The information is much worse at low slants than at high slants (the reasons for this are explored in some detail in ref. [16]). The subjects' performance qualitatively tracked the information content of the stimuli; It was quite poor at low slants, with thresholds for discriminating slants away from the fronto-parallel ranging from 29 to 46°. For slants of 70°, discrimination thresholds shrank to anywhere from 1.2 to 3.1°. To the extent that these thresholds reflect the subjects' ability to use texture information for perceiving surface slant in the natural world, it is clear that texture only becomes a useful cue at slants greater than 50° (to be generous). To give this some context, consider that a point on a horizontal ground plane approximately one eye-height away from an observer (1.75 m), has a slant of 45° relative to the local line of sight. One would have to look 2.75 eye-heights (4.8 m) away to see a point at 70° slant.

4.3.2. Why so inefficient?

Compared to the reliability of the information in the stimuli, subjects performed with quite low efficiency. Besides peripheral, image coding noise and central decision uncertainty, several 'perceptual' sources of inefficiency exist for a problem like slant discrimination form texture. We will briefly discuss four of the more interesting ones.

4.3.2.1. Limits on spatial integration. The performance of the ideal observers reflects optimal use of all of the

Fig. 2. Example stimuli from experiment 1. Subjects were asked to judge which of pair of stimulus surfaces like these had the greater slant. (a) The left surface is fronto-parallel (slant = 0°), the right surfaces has a slant of 50°. (b) The left stimulus has a slant of 73° and the right one has a slant of 60°. Both of these differences in slant were greater than the measured thresholds for subjects. Note the change in texel size which results from the re-scaling of the texture to normalize the area contrast in the images. The texels at the bottom of the image are larger (and the ones at the top, smaller) in the image of the more slanted surface. This reflects the difference in the scaling gradient between the two images.

information presented in a display. In experiment 1, this extended over a fairly large field of view (20° × 25°). To the extent that subjects did not efficiently integrate texture information over space, their efficiency will have been reduced.

4.3.2.2. Inefficient weighting of texture cues. To the extent that the visual system uses a sub-optimal weighting of cues, observers will be inefficient at discriminating slant-from-texture. For example, were subjects to have relied primarily on scaling information, they would have been inefficient relative to the total available information, even though their efficiency for using scaling information would have been high.

4.3.2.3. Incorrect prior assumptions. The subjects may have relied on incorrect prior assumptions about the statistical structure of surface textures, which could have led to a reduction in performance relative to an ideal observer which knows the correct structure. If the subjects did not assume isotropy, they would have been inefficient discriminators in this experiment—note that the thresholds for the foreshortening ideal observer which does not use prior knowledge of isotropy are almost as high as the human observers' thresholds.

4.3.2.4. Other slant cues and biases. Non-texture cues in the stimuli (lack of gradients in image blur which would be consistent with slanted surfaces, small amounts of motion parallax, etc.) may have affected discrimination performance. In these experiments, the non-texture cues should have suggested fronto-parallel surfaces. Since these were presumably fixed for all of the stimuli in the experiment, the cues would generally have made discrimination performance worse.

The next three experiments explored the effects on discrimination performance of perturbing a number of stimulus parameters. In experiments 2 and 3, we looked at the effects of changing the field of view dimensions (height and width of stimulus surfaces). In experiment 4, we looked at the effects of varying the reliability of scaling and foreshortening information. The results of these experiments shed some light on each of the perceptual issues discussed above.

5. Experiment 2

Blake et al. [20] were the first to systematically measure how texture cue reliability changed as a function of the size of one's field of view on a surface. In their ideal observer simulations, they fixed the number of texels in an image while varying the field of view symmetrically around the straightahead line of sight. The reliability of both texture cues studied (position and foreshortening of line elements) increased with increasing field of view. The major purpose of Blake et al.'s work was to show that the relative reliability of the density and foreshortening cues varied with field of view size, with density information dominating for large fields of view and foreshortening information dominating for small fields of view[3]. They were not particularly concerned with the first-order effects of field of view on cue reliability. It is with these first-order effects that we are concerned in the following two experiments.

As described in ref. [16], the dependence of texture information on field of view size is different in the direction of surface tilt (vertical for ground planes) than it is in the orthogonal direction. Changes in field of view size in the direction orthogonal to surface tilt have relatively simple effects on the informativeness of textures. To a first approximation, doubling the field of view on a surface in a direction orthogonal to the surface's tilt will have the effect of simply doubling the information content of the stimulus, by doubling the amount of visible texture[4]. Changing the field of view in the same direction of the surface tilt, on the other hand, has a considerably more complicated effect on the information content of a stimulus image. We will take up this point further in experiment 3. In this experiment, we look at how subjects discrimination performance varies with changes in the horizontal field of view on vertically slanted surfaces. Changes in discrimination performance with changes in horizontal field of view size should primarily reflect limits on the range of spatial integration of texture information (as opposed to strategic focusing of attention on selected stimulus regions).

5.1. Methods

The methods were essentially the same as those used in experiment 1. Psychometric functions were measured for textured surfaces with a test slant of 65° in four field of view conditions. These corresponded to horizontal fields of view of 2.5, 5, 10 and 20°. The vertical field of view was fixed at 25° as before. The experimental conditions were meant to simulate different size fields of view on the same set of texture patterns; therefore, the number of texels in a stimulus varied according to the field of view size, giving 19, 38, 75 and 150 texels per stimulus image for the four fields of view used. The differences in field of view were effectively created by changing the width of the windows bounding the texture stimuli. The inner edges of the windows remained fixed in all conditions (Fig. 4). In any given trial, test and target stimuli had the same field of view size. All

[3] The conclusions of Blake et al. should be tempered by the observation that relative cue reliability depends in large measure on the statistics of the ensemble of surface textures considered [16].

[4] The reliability of local texture information does vary somewhat with horizontal position in the image (for ground plane images), increasing slightly as one moves away from the direct line of sight. This increase is small near the straight-ahead line of sight, however.

other parameters for the stimuli were the same as in experiment 1. Fig. 4 shows examples of the stimulus pairs used in the experiment.

Four undergraduates naive to the purposes of the experiment and naive to vision science in general served as subjects for the experiment. All had normal or corrected to normal vision.

5.2. Results

In this and later experiments, in which the thresholds varies over a smaller range than in experiment 1, we would like some assurance that comparisons across the single threshold parameter of the psychometric model is a reasonable way to summarize the subjects' performance (i.e. that the results generalize over the entire range of the psychometric function). This consideration led us to apply a reduced psychometric model to the data. In the reduced model, the same variance stabilizing parameter, β, was assumed to apply to all of the experimental conditions—a condition which guarantees that psychometric functions with different thresholds will be non-overlapping, allowing us to use thresholds as unambiguous measures of relative performance in the different conditions. We tested the reduced model using a nested-hypothesis test [21] in which the relative likelihoods of the reduced and full models (with β allowed to vary between conditions);

$$2\log_e\left(\frac{L_{\text{independent}}}{L_{\text{simple}}}\right) \tag{5}$$

is distributed as χ^2 with, in this case, $8-5=3$ degrees of freedom. At $\alpha=0.01$, we could not reject the simplified model for any of the four subjects in the current experiment. This justifies our use of the reduced model for our analysis, and of thresholds as an appropriate measure of a subjects' discrimination performance.

Fig. 5 shows plots of the maximum likelihood estimates of subjects' 85% thresholds as a function of horizontal field of view size for all four subjects. In order to test for an effect of field of view on the subjects' thresholds, we ran a one-way ANOVA on the thresholds. The within-condition variances for the threshold estimates were derived using Fisher's approximation of the covariance matrix for the parameters of the reduced psychometric model (four thresholds and one stabilizing parameter)[5]. The results of the ANOVA are summarized in Table 1—the effect of horizontal field of view size is significant for all four subjects. Individual steps in thresholds were, by and large, too small to support strong pair-wise comparisons between neighboring thresholds; however, with the exception of subject HH, the data suggest that improvement in discrimination performance may extend over the whole range of field of view sizes tested.

5.3. Discussion

The subjects' thresholds decreased more slowly with the field of view size than would have been suggested by the decreases in ideal observers' thresholds; however, this could as easily have been a result of central sources of uncertainty as it was of inefficient spatial integration. The interesting result is that subjects' spatial integration appears to extend at least to 10° of visual angle in the horizontal direction and probably further than that for most of the subjects.

For the smallest field of view used, 2.5°, the thresholds of two of the subjects were lower than those of both the position ideal and the anisotropic foreshortening ideal. The other two subjects had thresholds which were close to these ideals'. Subject TT also beat the two ideals in the larger 5° field of view condition and matched the performance of the scaling ideal for the 2.5° condition. This provides strong evidence that the subjects were not relying only on the cues represented by those ideal observers. While it is possible, for example, that the subjects relied on a combination of these cues, they would have to have been very efficient in doing so to have achieved the level of performance that they did. The data, therefore, suggest that at least in the small field of view condition, subjects used some combination of the scaling and foreshortening-with-isotropy cues to make their discriminations. Furthermore, since no particular change in processing strategy is suggested for changing horizontal fields of view, we suspect that the conclusion generalizes to larger fields of view as well, though the data does not speak directly to this issue.

As noted in the introduction to this section, changes in field-of-view size in the direction of surface tilt have more complicated effects on the information content of textures than do orthogonal changes. The next experiment was designed to look at how such changes affect the human observers' performance.

6. Experiment 3

Blake et al.'s result that texture cue reliability increases with increases in the size of the field of view on a surface (for a fixed number of texels) pertains primarily to changes in the field of view size in the same

[5] The ANOVA assumes independence of the threshold estimates. In the reduced model, the threshold estimates become correlated through the shared induence of the variance stabilizing parameter on the thresholds. In this and all other experiments, the highest correlation was found to be $\rho=0.2$—low enough to have minimal impact on the F statistic.

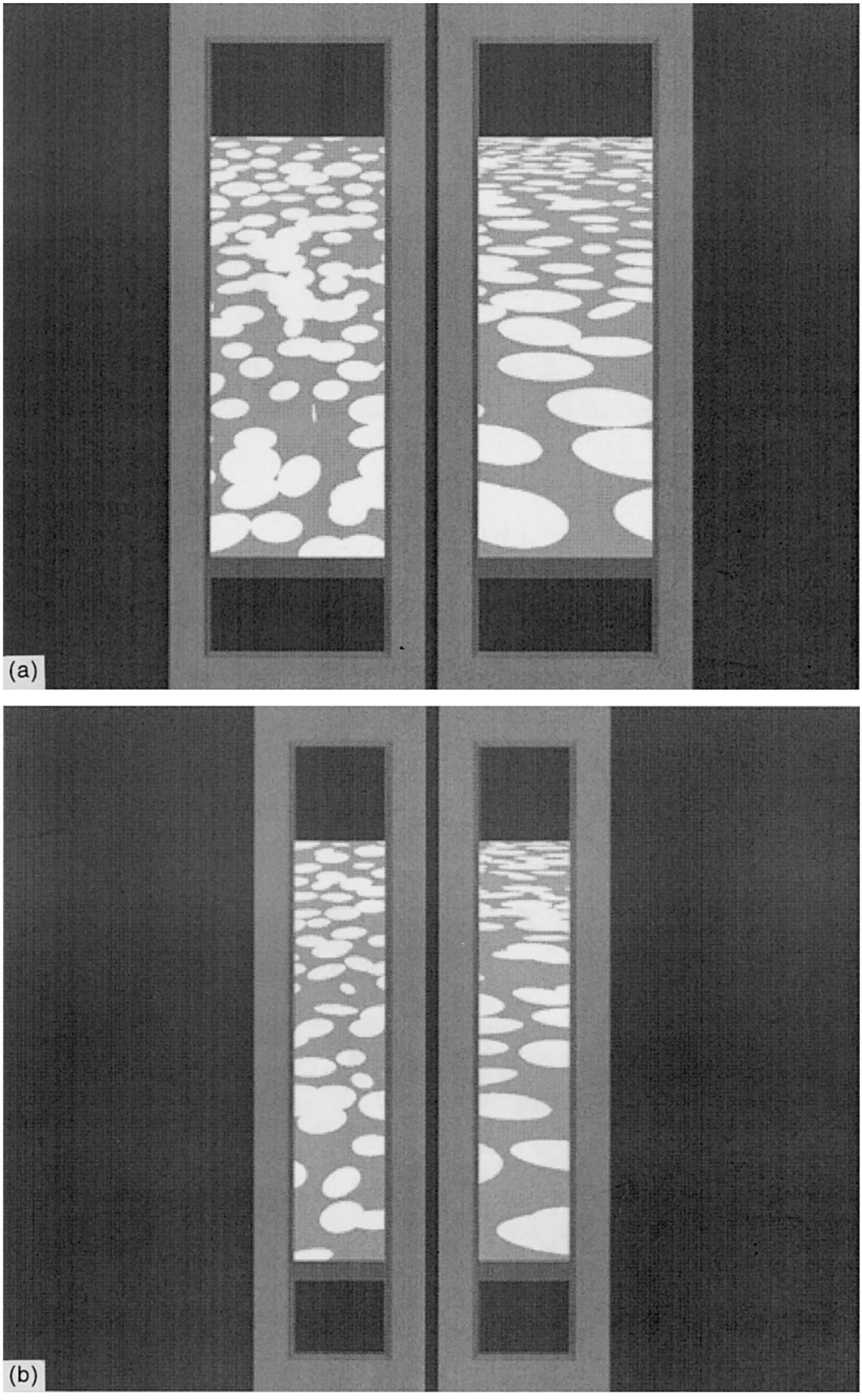


Fig. 4. Example stimuli for experiment 2. (a) Stimuli with a 10° horizontal field of view. (b) Stimuli with a 5° horizontal field of view. In both (a) and (b) the left surface has a slant of 60° away from the line of sight and the right surface has a slant of 73°.

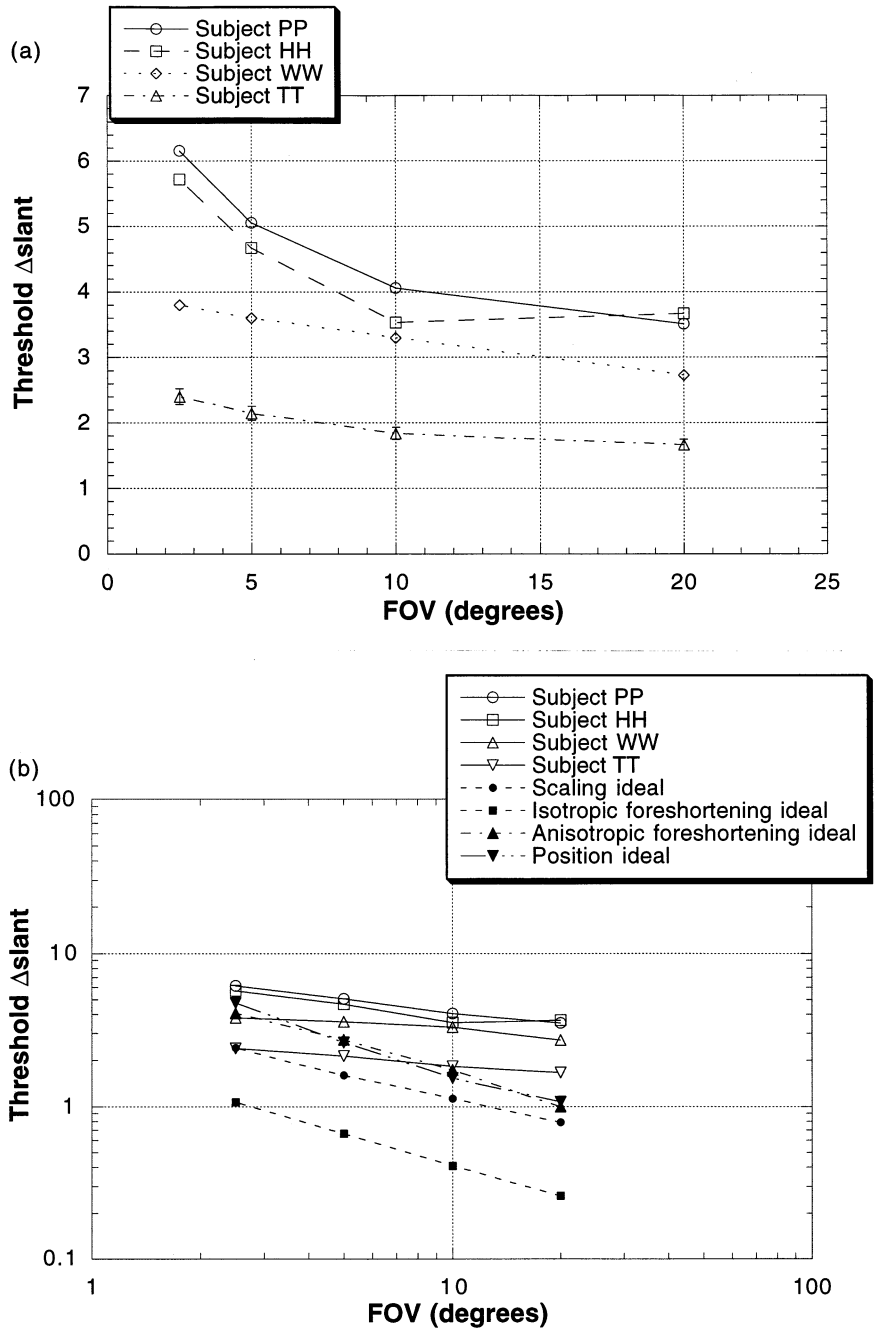


Fig. 5. (a) Linear plots of the subjects' thresholds as a function of horizontal field of view size in experiment 2. The error bars on the lower curve are shown to give an idea of the size of the standard errors in threshold estimates. (b) Log-log plots of the same thresholds shown for comparison with the thresholds of the four provisional ideal observers.

direction as a surface is slanted (in the tilt direction). we will henceforth refer to this direction as the vertical direction, as it is for the stimuli used in our experiments. For ease of explication, we will also assume a ground plane tilt (the top of the image corresponds to more distant regions of a surface). Three independent factors determine the effects of vertical field of view size on texture cue reliability: the relative density of texture elements in different parts of an image, the extent of

Table 1
F-scores from an analysis of variance done on the threshold data from experiment 2

Subject	$F(3,\infty)$	P value
PP	17.1	<0.001
HH	16.5	<0.001
WW	8.1	<0.001
TT	7.2	<0.001

Table 2
F-scores from an analysis of variance done on the threshold data for the four top views used in experiment 3 (including the largest field of view)

Subject	$F(3,\infty)$	P value
KK	9.4	<0.001
PP	0.8	**
YY	1.6	**
ZZ	2.0	**

spatial gradients contained within the image (relevant for the scaling, position and foreshortening-without-isotropy cues), and differences in the relative contributions of individual texels to texture cue informativeness as a function of position in the image. The interplay between these factors is discussed in detail in ref. [16]. We will only briefly review the main points here.

Density effects: Since the density of texture elements increases with distance of a surface away from the viewer, the top regions of an image will contain most of the texels in the image. This tends to increase the relative informativeness of texture in the top of an image.

Spatial extent of texture gradients: Texture scaling, position and foreshortening-without-isotropy cues are non-local, in that they depend for their informativeness on spatial gradients in local perspective distortions. Since decreasing the field of view reduces the total extent of such gradients, it will necessarily decrease the informativeness of these cues.

Position dependence of textel informativeness: For all of the texture cues, the relative contribution of individual texels to the information provided by the cues varies with position in the image—texels at the top of the image contribute relatively more to texture cue informativeness than do texels at the bottom. For the scaling cue this results from the greater scaling effect at larger distances, while for the foreshortening cue, it results from the fact that the local slant of a surface, relative to the line of sight, increases from bottom to top in the image, and small differences in slant lead to larger changes in texture foreshortening at high slants than at low slants. Since the position cue derives from both the scaling and foreshortening effects of perspective, it shows similar behavior to the scaling and foreshortening cues.

Table 3
F-scores from an analysis of variance done on the threshold data for the three botton views used in experiment 3 (including the largest field of view)

Subject	$F(2,\infty)$	P value
KK	79.9	<0.001
PP	152.6	<0.001
YY	326.3	<0.001
ZZ	101.1	<0.001

The end result of the factors described above is that 'bottom-up' reductions in field of view on a surface lead to significantly slower reductions in texture cue reliability than do 'top-down' reductions. In experiment 3, we measured the subjects' discrimination thresholds as a function of vertical field of view size for both bottom-up and top-down reductions in field of view size.

6.1. Methods

The methods were in all respects, except for the stimuli, the same as used in experiment 2. The field of view was controlled by adjusting the position of either the bottom or top edge of the stimulus surfaces in the image. Six fields of view were used: The largest was $20° \times 25°$, as in experiment 2. The bottom edge of the view was raised successively by halves to create three new fields of view with vertical extents of 12.5, 6.25 and 3.125°. We will refer to these as top views. Similarly, the top edge was lowered successively by halves to create two other fields of view with vertical extents of 12.5 and 6.25°, which we will refer to as bottom views. We did not use a 3.125° bottom view, since it would have contained only a few texels. The number of texels in the largest view was 150. The top views contained, respectively, 122, 86 and 54 texels. The bottom views contained 28 and 10 texels, respectively. Fig. 6 shows examples of the stimulus pairs used in the experiment. In any given trial, test and target stimuli had the same field of view parameters.

Four undergraduates naive to the purposes of the experiment and naive to vision science in general served as subjects for the experiment. All had normal or corrected to normal vision.

6.2. Results

We estimated the subjects' thresholds using the reduced psychometric model described in Section 5.2. A nested hypothesis test showed that we could not reject the reduced model for any of the subjects ($P < 0.01$). Fig. 7 shows plots of the subjects' thresholds as a function of the field of view size for both top and bottom views. We ran separate one-way ANOVAs for the bottom and top views. Within-condition variances for the threshold estimates were derived as in Section 5.2 (see also footnote 4). Both sets of views included the full $20° \times 25°$ view. For top views, only subject PP showed a significant effect of field of view. For bottom views, all subjects showed significant effects (Tables 2 and 3).

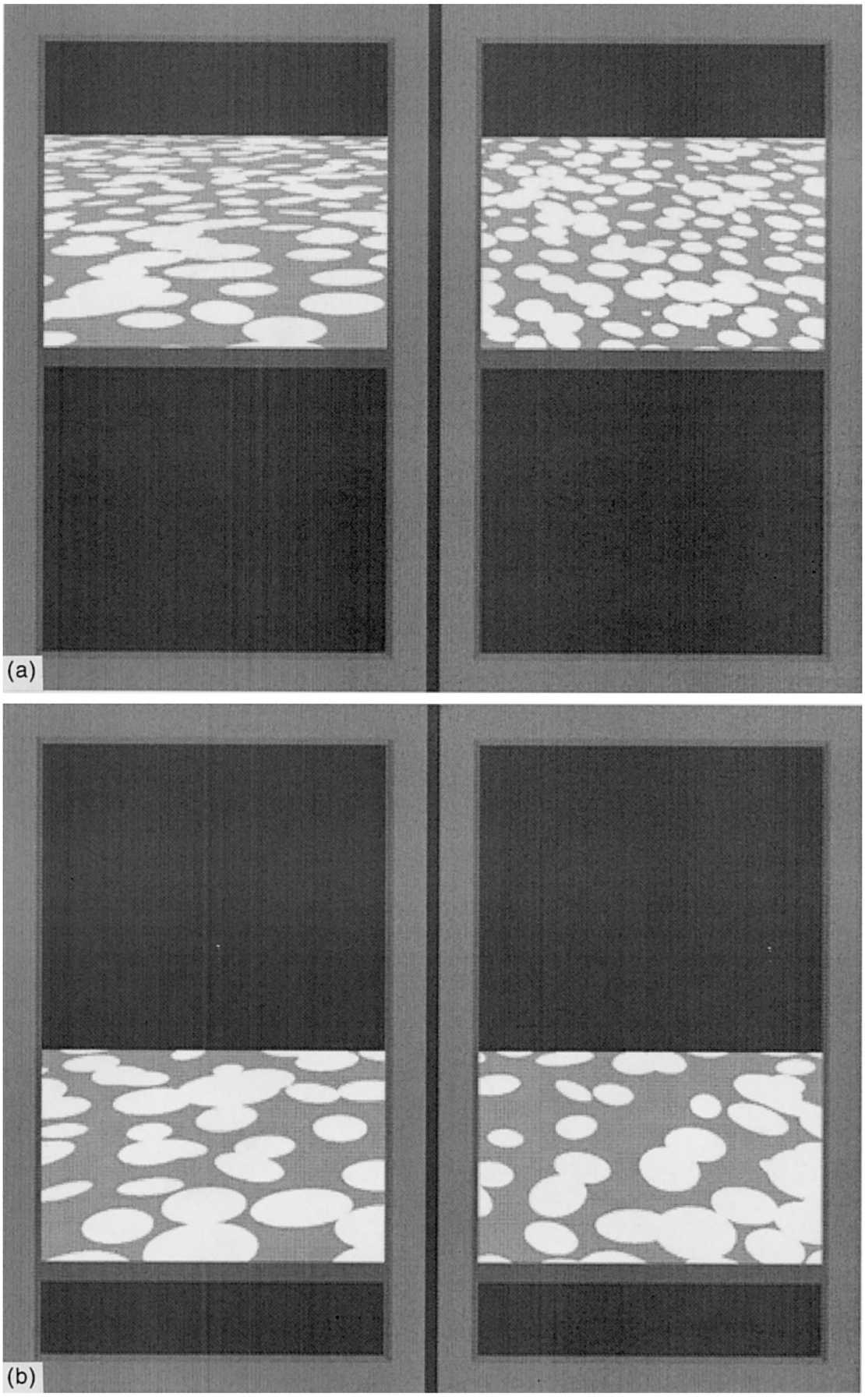


Fig. 6. Example stimuli for experiment 3. (a) 'Top' views of surfaces with the bottom half removed. (b) 'Bottom' views of surfaces, with the top half removed. Other conditions in the experiment were created by successively halving these views in either the bottom-up or top-down directions. In both (a) and (b) the left surface has a slant of 73° away from the line of sight and the right surface has a slant of 60°.

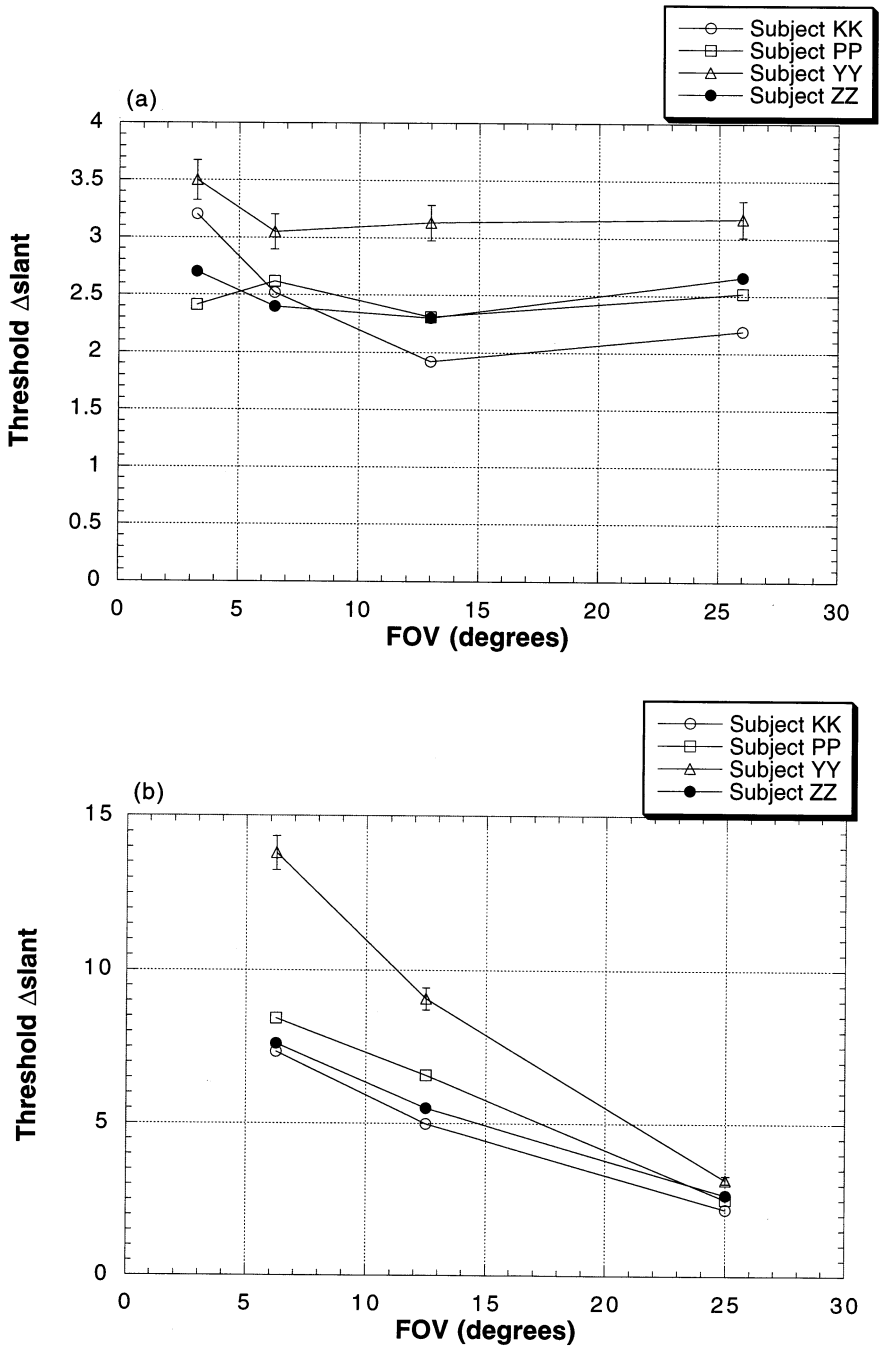


Fig. 7. Threshold curves as a function of vertical field of view size for experiment 3. (a) Thresholds for top views. (b) Thresholds for bottom views. The rightmost data point (25° field of view) in each graph is from the same experimental condition. Error bars are shown for one curve to illustrate the magnitude of the standard errors. (c) and (d) show log-log plots of the thresholds for both human subjects and the four provisional ideal observers as a function of field of view size for top and bottom conditions, respectively.

6.3. Discussion

Three of the four subjects showed no significant change in threshold levels as a function of field of view size for top views of the stimulus surfaces. This is to be contrasted with the significant effects obtained with changes in horizontal field of view size for all the subjects in experiment 2. The latter result suggests that the lack of effect obtained for bottom-up changes in field of view was not simply due to limits on the spatial range of integration. The implication is that the subjects' strategically focused on the top parts of surfaces to make their discriminations, a strategy which is, in general, consistent with the spatial structure of texture information.

Comparing the performance of human observers and

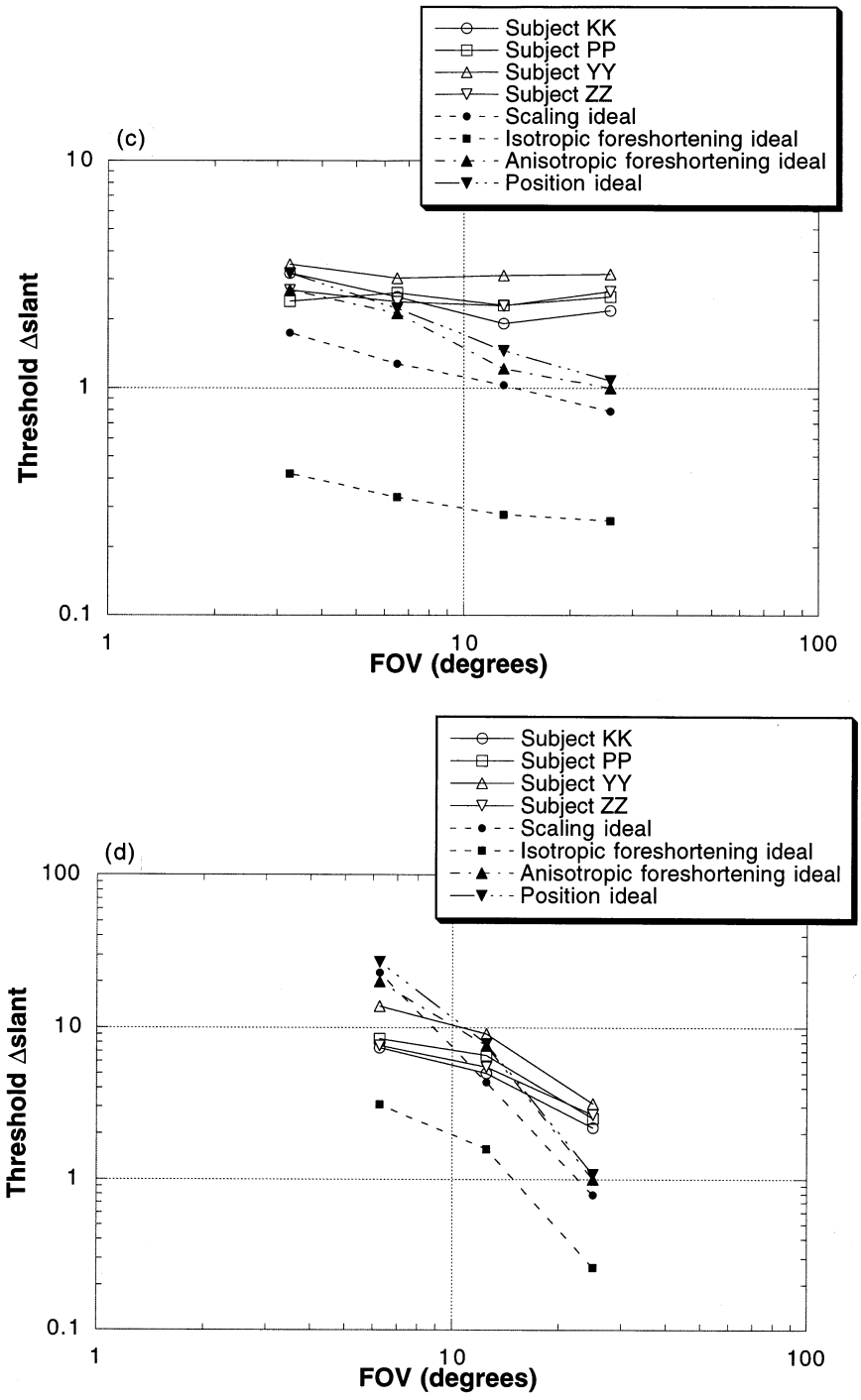


Fig. 7. (*Continued*)

the provisional ideal observers provides more interesting insights into how the subjects used the texture cues to make discriminations. In the smallest top view, all of the subjects had thresholds less than or equal to that of the foreshortening-without-isotropy ideal, and two of the subjects had thresholds less than or equal to that of the position ideal observer. Moreover, in this condition, the subjects' thresholds were no more than twice that of the scaling ideal. If the subjects relied solely on any of these three cues, one would expect to have found a change in threshold across the different field of view conditions for the top views. The data from the bottom views is even more striking. Thresholds for all four subjects were lower than those of the scaling, position and foreshortening-without-isotropy ideal observers in the smallest field of view condition. Moreover they were close to the thresholds of the three provisional ideals in the larger, 12.5° field of view condition. This clearly

indicates that the subjects relied primarily on the foreshortening cue in the bottom field of view conditions.

Taken together, the results of this experiment and the previous one are consistent with the hypothesis that subjects in many viewing conditions relied to a significant degree on foreshortening information to make slant discrimination judgments. Subjects may, however, have switched strategies for different field of view conditions. For example, they may have only been able to use the scaling cue with large vertical fields of view, even though the cue is theoretically of some use in smaller field of view conditions.

7. Experiment 4

In the fourth experiment, we measured subjects' discrimination thresholds for texture stimuli with varying degrees of cue reliability. In particular, we independently varied the reliability of the scaling and foreshortening cues in the stimuli. This was done by independently manipulating the variances of surface texel lengths (for scaling) and of surface texel shapes (for foreshortening). At the most superficial level of analysis, the results will help elucidate the contribution of central 'noise' to the subjects' performance—small changes in thresholds across large changes in texture reliability would suggest a large influence of central factors, while large changes in threshold would suggest a smaller influence of central factors. Of more consequence for general theories of slant from texture are the potential implications of the results for texture cue weighting. Significant decreases in thresholds with decreases in the reliability of one cue suggest that a subject makes use of that cue in their judgments. Besides looking for differential effects of scaling and foreshortening cue reliability on subjects' performance, we will also compare the subjects' performance with that of the provisional ideal observers over the different stimulus conditions to gain insight into how the subjects use the different texture cues to make their judgments.

Four stimulus conditions were used in the experiment, corresponding to mixtures of two levels of the scaling and foreshortening cue reliability factors—strong $\times$ weak. Of course, the distinctions between strong and weak cues are relative.

7.1. Methods

In order to maximize efficiency, while maintaining a reasonable field of view size, we limited the size of the stimulus display area to $10° \times 12.5°$, keeping the average density of texels in the image the same as before. Stimuli in the experiment, therefore, contained 60 texels each. Four classes of stimuli were generated based on a crossing of two different distributions of surface texel sizes and two different distributions of surface texel shapes. The two surface texel size distributions were Gaussians (truncated just above zero) with standard deviations either 5 or 40% of the mean. As described previously, the mean of the distributions was adjusted for each stimulus to ensure that a fixed percentage of image area was covered by texels. For this experiment, the coverage was set to 40% rather than the previously used 50%, to adjust for the increased likelihood of texel overlap which resulted from increasing the variance of texel properties. The distribution of texel aspect ratios was the same as before for one class of texel shapes, while for the other, more irregular class, it was approximately uniform between 0.2 and 1. The test slant for all conditions was set to 65°.

Fig. 8 shows examples of the most regular and most irregular textures used in the experiment.

Two non-parametric (four-up/one-down and one-up/four-down) staircases were interleaved for each condition to find the 85 and 15% threshold differences in surface slant needed for subjects to correctly judge which of a pair of test and target stimuli had greater slant. The procedure corresponds to searching for the 85% thresholds for discriminating surface slants both above and below the test slant. For the final data analysis, we computed the maximum likelihood estimates of psychometric function parameters to estimate thresholds, so the staircases served primarily to sample informative regions of the psychometric function. The subjects ran in eight blocks of trials each containing 50 steps of the staircases for all four conditions randomly interleaved, giving a total of 400 trials per session. The first session was discarded as practice.

Four undergraduates naive to the purposes of the experiment and naive to vision science in general served as subjects for the experiment. All had normal or corrected to normal vision.

7.2. Results

We estimated subjects' thresholds using the reduced psychometric model described in Section 5.2. The nested hypothesis test showed that we could not reject the reduced model for any of the subjects ($P < 0.01$). Fig. 9 shows plots of each of the four subjects' thresholds across the stimulus conditions used in the experiment. To measure the significance of the effects of texture cue reliability, we performed two-way analyses of variance on the estimated thresholds for each subject. We used the average variance of threshold estimates derived from the model fitting procedure as

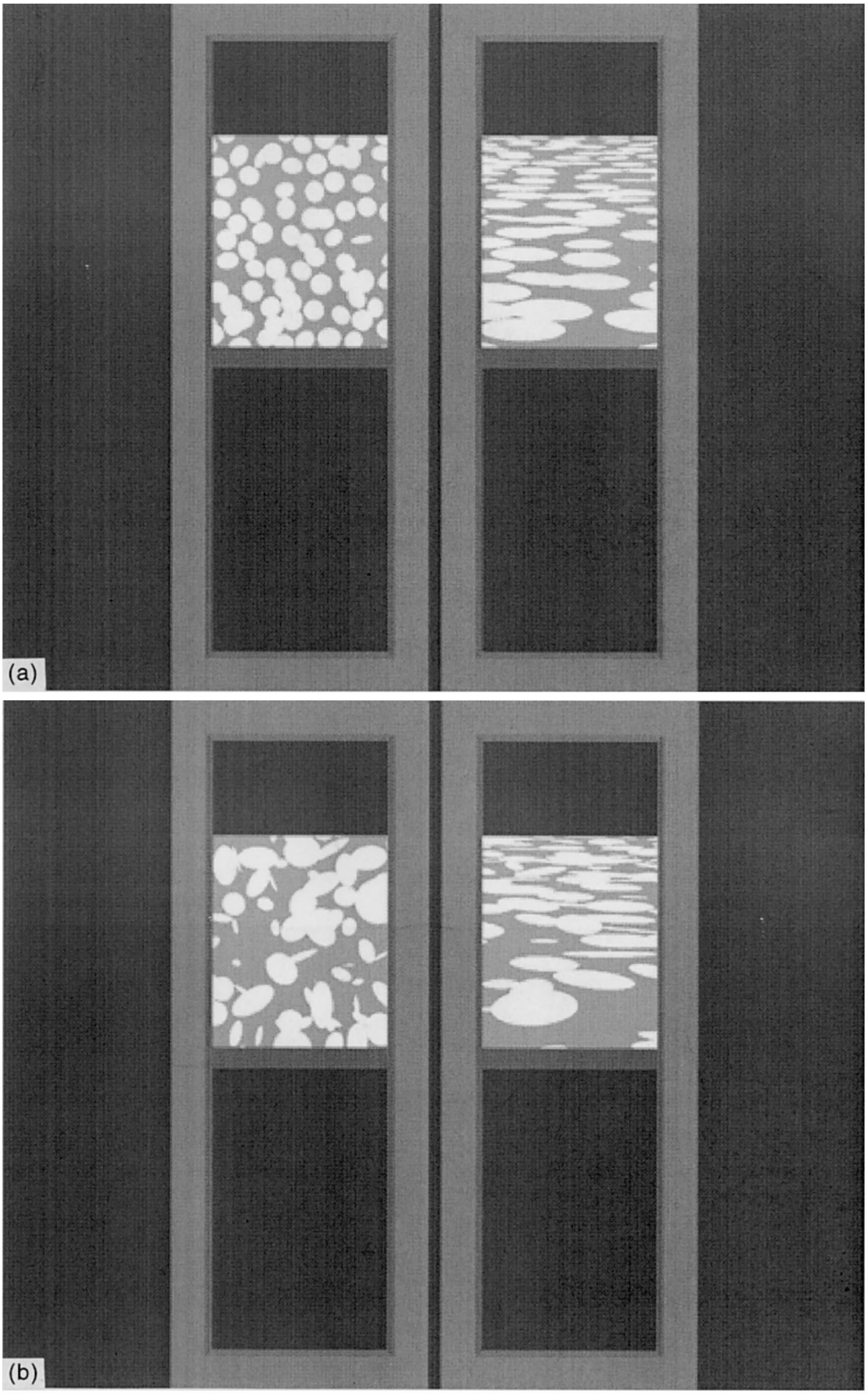


Fig. 8. Example stimuli for experiment 4. (a) Textures with both regular texel sizes and regular texel shapes. (b) Textures with both irregular texel sizes and irregular texel shapes. In both (a) and (b) the left surface is fronto-parallel, to illustrate what the textures look like on the surfaces, and the right surface is slanted away from the line of sight at 73°.

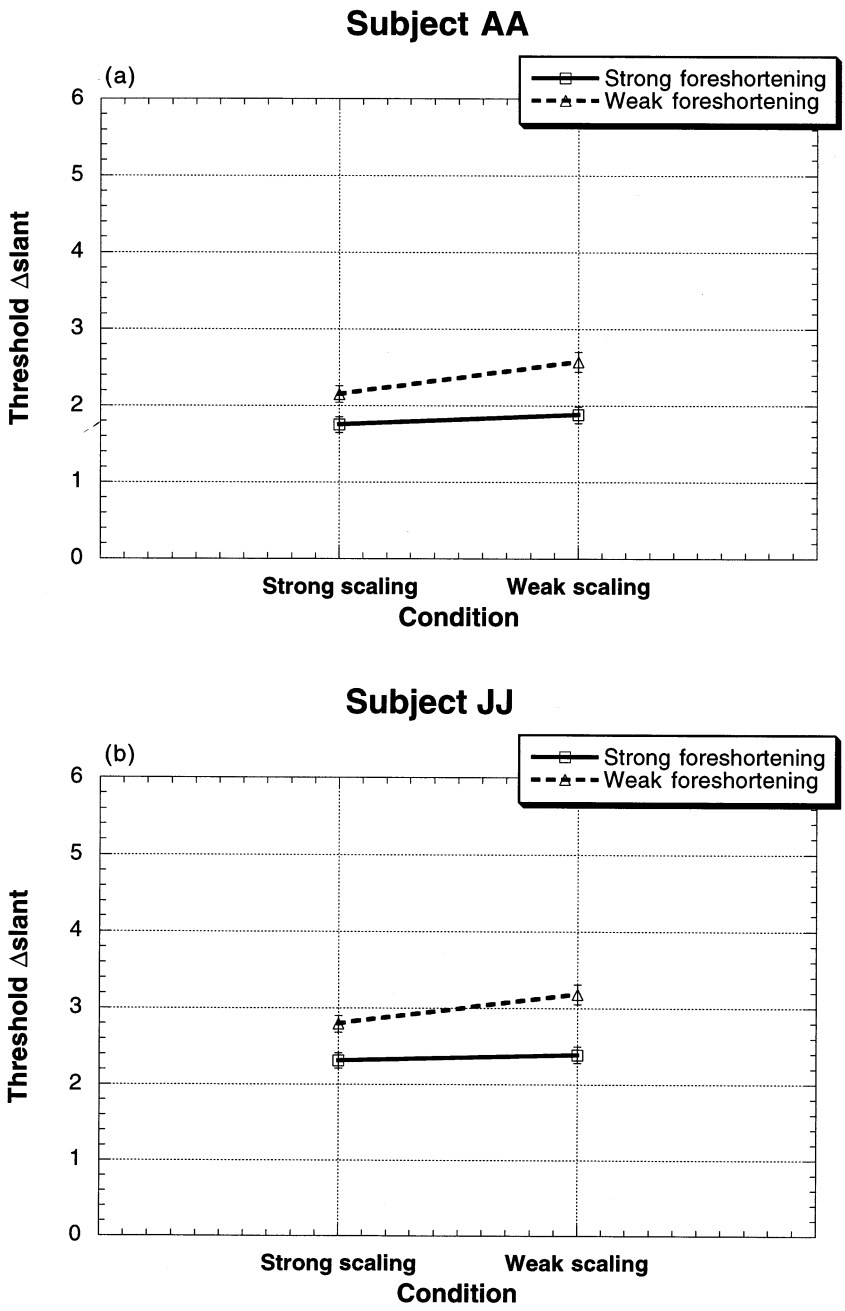
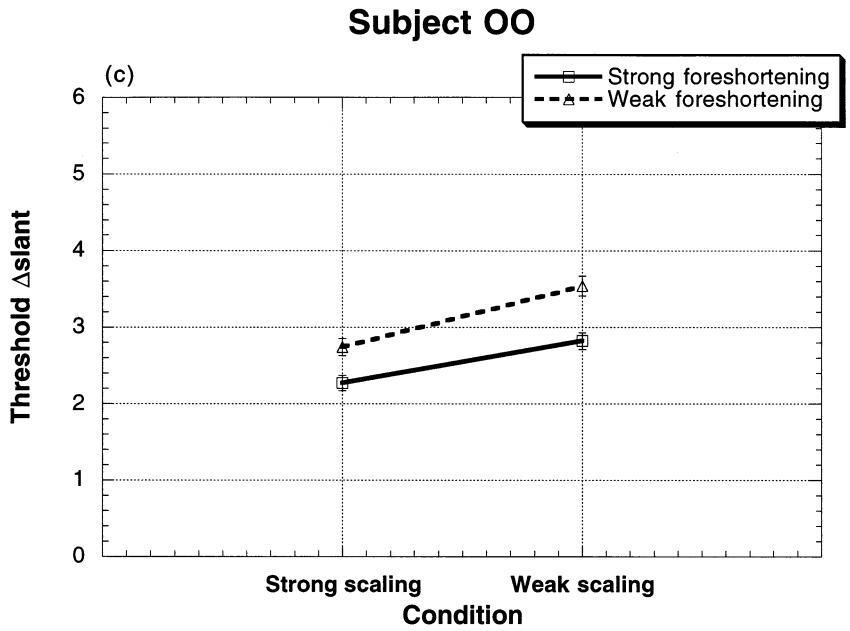


Fig. 9. Thresholds for each of the four subjects for the four experimental conditions of experiment 4. The thresholds plotted here are for the positive difference in slant needed to correctly judge (85% of the time) that a target surface is more slanted than a test surface at 65° slant.

the within-condition variance for the ANOVA. The results are tabulated in Table 4. Subjects AA and JJ showed only main effects of foreshortening cue reliability, while subject OO showed main effects of both foreshortening and scaling cue reliability. No significant interactions were found for these three subjects. Subject ZZ showed no significant main effects, but did show a significant interaction. This resulted from the somewhat paradoxical behavior of having best thresholds when only one of the two cues was strong.

7.3. Discussion

We would like to use the pattern of results across different cue reliability conditions to draw inferences about the relative weights assigned by subjects to different cues. Doing so would lead to the conclusion that subjects AA and JJ relied primarily on foreshortening information and subject OO relies to some significant degree on both cues. Such conclusions should be tempered by the fact that changes in the threshold across different conditions was small relative to the changes in



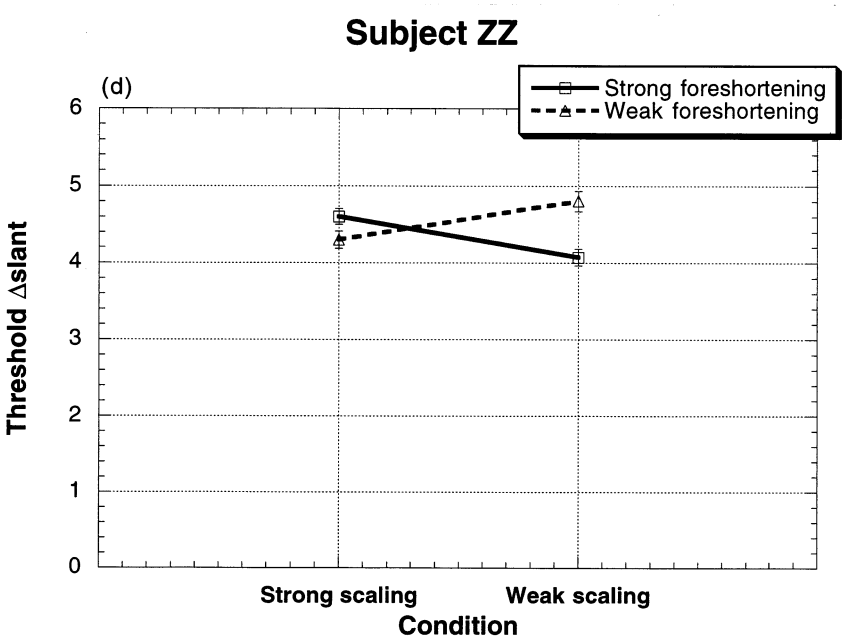


Fig. 9. (*Continued*)

cue reliability induced by the stimulus manipulations (Fig. 10). Moreover, we should note that manipulations of one cue might have affected the subjects' ability to use another of the cues. For example, increasing the variance of surface texel shapes would decrease the performance of an observer who relied on gradients in the horizontal width of texels, a sub-optimal form of the scaling cue (because it would increase the variance of the width of surface texels).

Table 4
F-scores from a two-way analysis of variance done on the threshold data from experiment 4

Subject	$F_{scaling}$ $(1,\infty)$	$F_{foreshortening}$ $(1,\infty)$	$F_{interaction}$ $(1,\infty)$
AA	3.3 (**)	12.3 ($P<0.001$)	19 (**)
JJ	2.3 (**)	16.1 ($P<0.001$)	2.3 (**)
OO	10.2 ($P<0.01$)	8.2 ($P<0.01$)	0.9 (**)
ZZ	0 (**)	1.0 (**)	8.3 ($P<0.01$)

Comparisons between the subjects' thresholds and the provisional ideal observers' thresholds suggest an alternative route to interpreting the data from the experiment. Fig. 10 shows the subjects' thresholds plotted with the thresholds of each of the four provisional ideal observers. The comparisons allow us to reject a number of single-cue theories of how the subjects' performed the discriminat task.

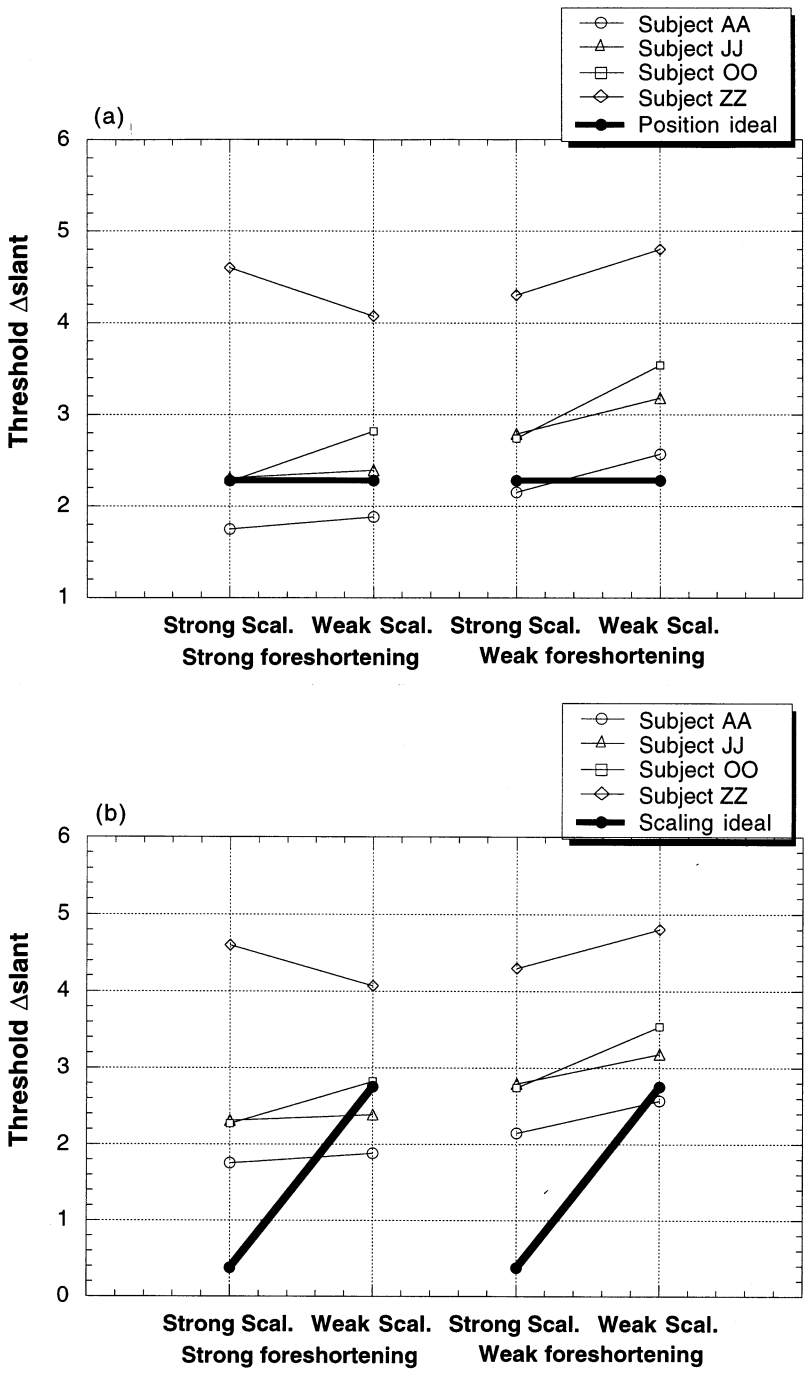


Fig. 10. Subjects' thresholds from experiment 4 replotted for comparison with each of the four provisional ideal observers. In the figure legends, 'fore.' refers to foreshortening information and 'scal.' refers to scaling information. (a) The subjects' thresholds compared against those of the position ideal observer. Thresholds are plotted as a function of the scaling cue reliability, with one plot each for strong and weak foreshortening conditions. (b) The subjects' thresholds compared against those of the scaling ideal observer, plotted in the same way as (a). (c) The subjects' thresholds compared against those of the foreshortening ideal observer which does not assume isotropy. Thresholds are plotted as a function of the foreshortening cue reliability, with one plot each for strong and weak scaling conditions. (d) The subjects' thresholds compared against those of the foreshortening ideal observer which assume isotropy, plotted in the same way as (c).

The position-only model: Fig. 10a shows subjects' thresholds as compared with those of the position ideal observer. Thresholds for three of the four subjects equaled or bettered that of the position ideal observer for stimuli with the most reliable scaling and foreshortening information. This definitively shows that these subjects did not rely only on the position cue to make judgements in that condition (Blake et al. used the same logic in testing a density-only model for curvature estimation [12]).

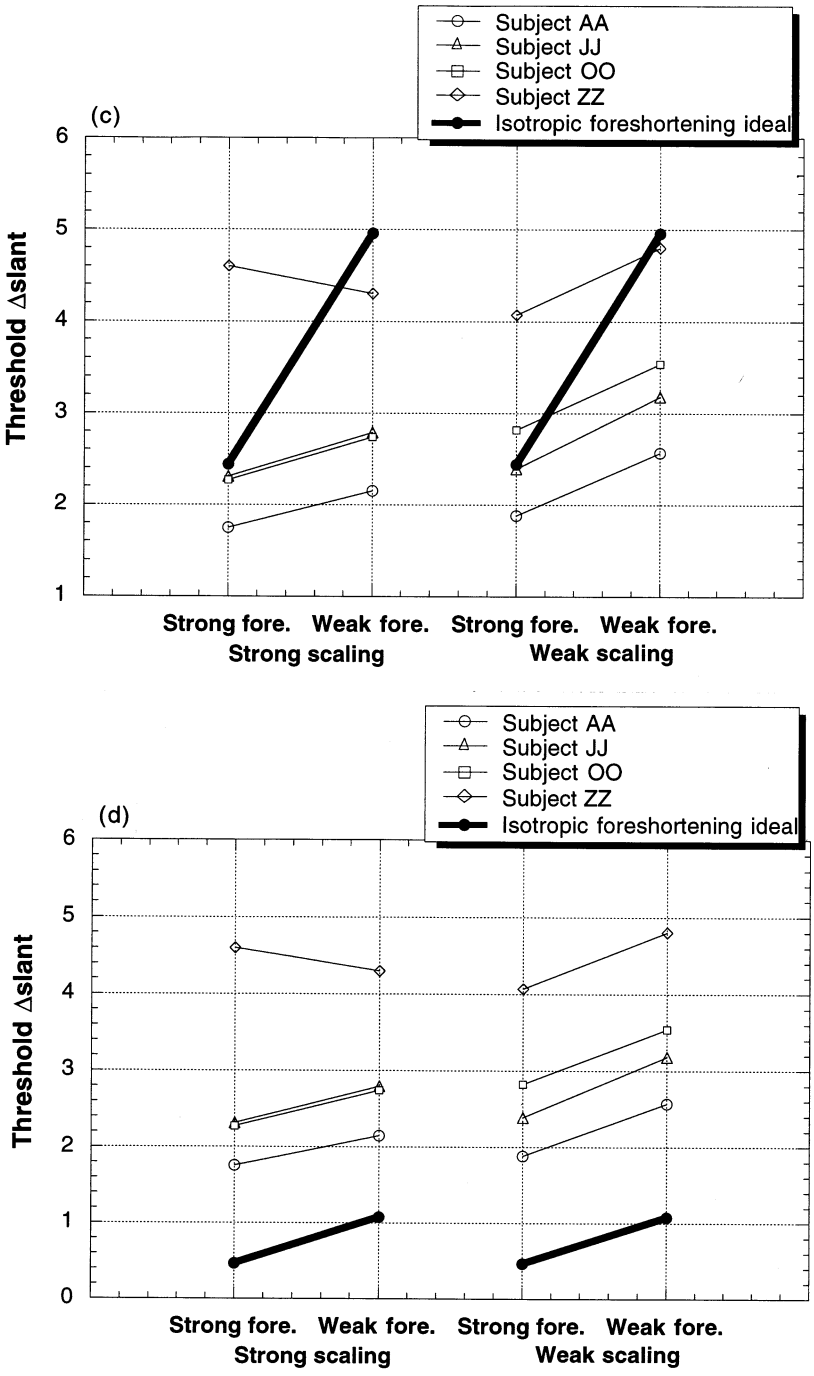


Fig. 10. (*Continued*)

The scaling-only model: Fig. 10b shows the subjects' thresholds as compared with those of the scaling ideal observer. As with position information, a condition can be found in which three of the subjects equaled or bettered the scaling ideal observer, indicating that scaling could not have been the only cue used by those subjects to perform the task in that condition. Moreover, were the subjects to have given a significant weight to the scaling cue, one would reasonably expect a larger effect of decreasing the reliability of the cue than is found. To see this, note that the thresholds obtain in the weak scaling/strong foreshortening condition, a large weight given to scaling would require highly efficient use of the cue (low internal noise). This would imply that observers should be highly sensitive to the reliability of scaling information. The data shows this not to be the case, suggesting that in at least the conditions in which scaling was weak, subjects relied heavily on other cues.

The foreshortening cue (without isotropy): Fig. 10c shows the subjects' thresholds as compared with those of the foreshortening ideal observer which does not use prior knowledge of surface texture isotropy. All observers outperformed this ideal observer in the weak foreshortening conditions and three of the subjects approached or bettered it in the strong foreshortening condition. The comparison clearly eliminates foreshortening without isotropy as a significant cue for subjects in this experiment.

The foreshortening cue (with isotropy): Fig. 10d shows the subjects' thresholds as compared with those of the foreshortening ideal observer which assumes isotropy. Clearly subjects perform much worse than the foreshortening ideal. It is intriguing to note that three of the subjects have threshold curves which are approximately parallel to the threshold curve for the foreshortening ideal observer (particularly as compared with the other ideal observers). This is consistent with the hypothesis that, for these three observers, the foreshortening cue was the dominant source of information for making discriminations (e.g. with an approximately constant level of internal noise). It does not however, prove the point.

The comparisons described above strongly point to the conclusion that foreshortening information contributed significantly to the subjects discrimination performance. The data from the current experiment, however, can only be used to draw qualitative conclusions about the subjects' cue integration strategies. Cue perturbation experiments [24] would be required to quantify these strategies.

Subject ZZ deserves some brief comment. Her thresholds were uniformly high, rendering the above analysis of little use for her data. She also showed the seemingly paradoxical behavior of performing better when only one of the two texture cues was strong than when both were strong. This suggests the possibility that the subject switched to an inefficient strategy on some conditions. One such possibility would be a switch from using foreshortening information to using scaling information for the most regular textures. Were the subject less efficient in her use of scaling information, this would represent a sub-optimal strategy and could lead to poor performance when the texture was in fact most regular. Of course, conclusions such as this are speculative at best, but they do illustrate the possibility of sub-optimal selection of cues.

8. Experiment 5

In experiment 5, we tested whether the results obtained using the elliptical element textures would generalize to more natural textures. In order to do this, we needed to create a more naturalistic class of textures than the elliptical element textures, but one with approximately the same statistical structure. We chose to use for natural textures a type of texture introduced into studies of shape-from-texture by Rosenholtz and Malik [15]. These are textures composed of so-called Voronoi polygons, generated from a random lattice of sample positions by tiling the surface with polygons that bound regions containing points which are nearer to one or another of the sample positions in the lattice. Shrinking the resulting polygons around their center of mass results in the type of texture pattern shown in Fig. 1b.

We measured subjects' discrimination thresholds for an ensemble of Voronoi textures and an ensemble of elliptical element textures with equivalent first-order statistics. We also measured points of subjective equality between Voronoi and elliptical element textures to determine whether the Voronoi textures elicited stronger percepts of surface slant, as might be suggested by their greater realism.

8.1. Methods

To create a set of Voronoi textures for the experiment, we first generated random lattices in exactly the same way as we created texel positions for the elliptical texture element stimuli. We then used a canned program[6] [25] for determining the Voronoi polygons for each lattice. We shrunk the polygons around their center of mass by a factor of 80%, to generate textures which appear like a tiled road. Subjects all report that these textures appear more realistic than do the elliptical element textures.

In order to control for the information content of the different types of texture patterns, we measured the first-order statistics of the second-order spatial moments of the Voronoi textures, in a procedure that amounted to fitting each polygon in the pattern with an ellipse and computing statistics of the resulting ellipse parameters (lengths, aspect ratios and orientations). We used these to generate elliptical element textures whose statistics were matched to the derived first-order statistics of the Voronoi textures.

Stimuli subtended a $10° \times 12.5°$ field of view and contained 60 texels each. Non-parametric staircases were run in three stimulus conditions: elliptical element versus elliptical element textures, Voronoi versus Voronoi textures and elliptical element versus Voronoi textures. In the last condition, the Voronoi textures served as test stimuli, and had a fixed slant of 65°. Three staircases were interleaved for each condition: four-up/one-down, one-up/one-down and one-up/four-down. The procedure was designed in part to minimize the possibility, in the crossed condition, of subjects basing their judgments solely on the type of texture pattern in a stimulus, since it insured that subjects

[6] The code is available by anonymous ftp from netlib.att.com.

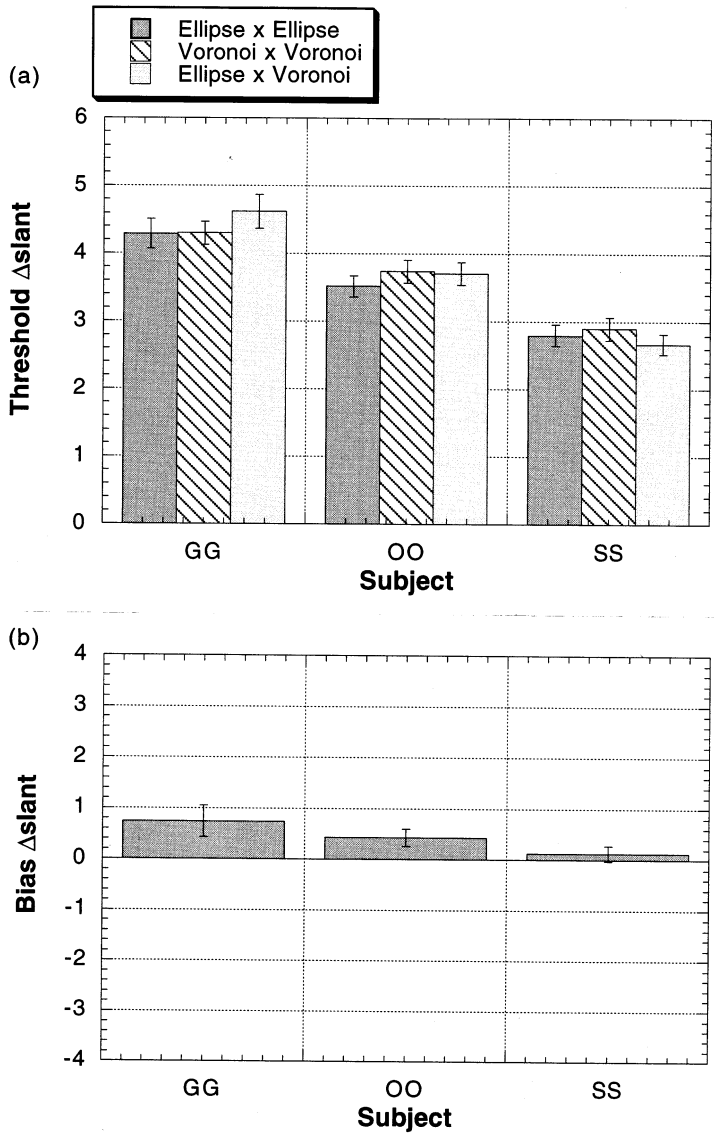


Fig. 11. (a) Threshold data for three subjects as a function of experimental condition in experiment 5. (b) Estimates of the difference in slant away from a test slant of 65° for the subjects to see an elliptical element target texture as having the same slant as a Voronoi texture. The bias provides an estimate of the difference in perceived slant one can expect to find between Voronoi and elliptical element textures with physically equal slants. The positive biases indicate that, if anything, the subjects see Voronoi textures as slightly more slanted than elliptical element textures.

selected stimuli from each class in approximately equal proportions over the course of the experiment.

Three undergraduates naive to the purposes of the experiment and naive to vision science in general served as subjects for the experiment. All had normal or corrected to normal vision.

8.2. Results

We estimated subjects' thresholds using the reduced psychometric model described in Section 5.2, with the addition of a free parameter for the 50% point of the psychometric function for the crossed stimulus condition. The nested hypothesis test showed that we could not reject the reduced model for any of the subjects ($P < 0.01$). Fig. 11a shows the 85% thresholds for the three subjects. Fig. 11b shows the difference in slant between stimuli needed to match the slants of surfaces with elliptical element textures to the 65° slant of the test Voronoi texture. An analysis of variance on thresholds using within-condition variances estimated from the psychometric function fit showed no significant effects of texture condition on the subjects' discrimination thresholds. T-tests on the points of subjective equality for condition three did show significant differences from the test slant of 65° for two subjects, GG and 00 ($P < 0.05$). The effects, however, were marginal ($< 1°$).

8.3. Discussion

The results support the generalization of previous results obtained with elliptical element textures to other, more naturalistic textures. Despite the fact that subjects noted the more natural appearance of the surfaces with Voronoi textures, the discriminability of the textures was not improved over that of the elliptical element textures, nor did they appear much more slanted than the elliptical element textures.

9. General discussion

The results of the experiments answer four general questions: how reliably does texture determine percepts of surface slant at different slants; how well does the visual system integrate texture information over the extent of an image; which texture cues does the visual system rely on to make judgments about surface slant; and does the visual system incorporate an assumption of isotropy? Experiments 1, 2 and 3 directly dealt with the first two of these questions. Answers to the other two questions can be indirectly inferred from the pattern of subjects' performance in these experiments and in experiment 4.

9.1. The effect of surface slant

The results from experiment 1 clearly show that the texture patterns used in this study were only effective cues for slant discrimination at large slants. Moreover, subjects' ability to discriminate surface slant from texture improved monotonically with increasing slant. This is entirely consistent with the ideal observer analysis, which shows that the informativeness of texture information about surface slant improves dramatically with increasing surface slant. These results are important for studies of texture cue interactions with other cues. They suggest that texture information may play a significant perceptual role at high slants (e.g. for ground planes under normal viewing), but much less so for small slants. The result can be potentially generalized to curved surfaces, where the increased informativeness of texture information in regions of high slant suggests that texture cues for curvature will be much more informative there. Interestingly, several studies of shape from texture use views of surfaces which minimize the potential of texture information, by showing head-on views of cylinders [4,11,12]. We suspect that larger perceptual effects of texture information will be found for other views, in which a curved surface is slanted away from the viewer.

9.2. Spatial integration

The results of experiment 2 clearly show that humans can integrate texture information over fairly large regions of an image. When roughly equivalent forms of information are added to a stimulus by increasing the field of view on a surface in a direction perpendicular to surface tilt, the subjects' discrimination performance improves. Changing the field of view in the direction of surface tilt leads to different effects on the subjects' performance. The data suggests a strategy in which the subjects gave preferential weighting to texture information in the top part of the displays. This is consistent with the results of the ideal observer analyses which show that the texture information in the top part of displays of ground surfaces (more generally, the more 'distant' parts of images) contributes disproportionately to the total information content of texture patterns. Thus, the subjects' performance may primarily reflect the informational structure of the stimuli, rather than a particular focusing of attention in selected regions of the displays. Of course, the analysis also suggests that if attention must be limited spatially, it should be directed to the tops of the displays.

9.3. Weighting of different texture cues

We have not explicitly attempted to measure the perceptual weighting of different texture cues. Our results do, however, provide some clues to how the visual system weights the different cues. We will limit our discussion to the foreshortening and scaling cues. No studies to date have found significant effects of texture density as a cue to surface orientation or curvature [7,8,11,13,20]. More generally, gradients in texel area (an analogue of density) seem to contribute little to subjects' performance [9,11]. The fact that subjects can equal or beat the position ideal observer in several conditions of the experiment is consistent with previous results. In general, we feel that position information in stimuli like ours is not of much practical use to the visual system.

The results of experiment 3 show that foreshortening information is used at least in degenerate conditions, such as the small-angle bottom views used there. This is hardly surprising given the theoretical unreliability of the other cues in such conditions. More relevant to the general problem is the pattern of results across different stimulus conditions. In experiment 3, we found that in the small-angle top views, the subjects approached or bettered the performance of the scaling ideal observer. Similar results were found in experiment 4 with larger fields of view (12.5° vertically) when the scaling cue was made less reliable by increasing the randomness of texel sizes on stimulus surfaces. In both experiments, subjects' thresholds were close to that of the scaling ideal

observer in the conditions in which scaling information was least reliable (sometimes beating it). Were the subjects to have relied heavily on the scaling cue, the results would imply great efficiency in using the cue. This, in turn, would imply that their performance should have improved significantly when the reliability of the scaling cue was increased (by increasing the field of view in experiment 3 and reducing texel size randomness in experiment 4). The lack of such an effect strongly suggests that subjects gave a significant weight to foreshortening information for making their discriminations. It does not, however, suggest the opposite, that scaling information is not used.

Researchers who have directly tested the relative contributions of foreshortening and scaling information have obtained conflicting results. In apparent conflict with our conclusions, Cutting and Millard's [8] work suggests the minimal importance of foreshortening information for the perception of planar surfaces. This result, however, was obtained using a task in which the subjects were asked to judge the planarity of surfaces, a task which did not directly measure orientation perception. In fact, it is unclear what combination of perceptual factors underlay the subjects' performance in their experiment. It may have included elements of orientation and curvature perception as well as less quantifiable properties like surface goodness (the degree to which stimuli appeared like solid surfaces), which varied in our experiments between elliptical element textures and Voronoi textures. Moreover, they relied on gross cue conflicts which may lead to observer strategies which do not reflect normal perceptual processing [24].

The results of a series of studies by Frisby and Buckley are more relevant to the subject's use of texture information to discriminate surface slant. These studies measured perceptual interactions between texture and stereo information in the estimation of the surface orientation. One study [10], using regular texture patterns (circles and rectangles), showed that both foreshortening (which they refered to as compression) and scaling information were used by the subjects to estimate slant, with either equal weight given to both cues or more weight given to foreshortening, depending on viewing conditions. Another study [14] found similar results, but with scaling being given more weight for slant judgments. The authors point out, however, that linear perspective, a cue which does not appear in the stochastic stimuli used here, may have provided the salient cue in their stimuli. In neither of these studies did the authors independently manipulate the shapes of texture elements and their sizes (what we refer to as foreshortening and scaling information). Rather, they independently manipulated their height and width measured in a coordinate frame aligned with surface tilt. This had the effect that changes to the scaling cue (the widths of texture elements), independent of changes to what they call the compression cue, also induced spatial distortions in the shapes of texels. What impact this would have on their results is unclear.

In a third study, Buckley et al. [13] used stochastic textures and looked for significant effects on slant estimates of making foreshortening, scaling and density cues inconsistent with stereo information. They found a significant effect of foreshortening information but not of either density or scaling. The result suggests greater reliance on foreshortening information, though their data plots suggest some effects of scaling, even if they were not large enough to be significant in their design. It is not clear from their description how they manipulated the scaling and foreshortening cues. If they used the same technique as in the previous study, their results may have been similarly confounded.

Our results are broadly consistent with those of Frisby et al. [14] and Buckley et al., [13] though we feel their results should be treated with some caution because of the methodological concerns described above. If anything, the current results suggest that subjects relied most heavily on foreshortening information, though the results do not support computing actual cue weights. Why should the visual system rely strongly on foreshortening information to perceive surface orientation from texture? Two reasons immediately present themselves upon looking at the ideal observer analyses presented here and in the companion article [16]. First, under the assumption of isotropy, foreshortening information is highly reliable across a broad range of surface texture ensembles. Second, the foreshortening cue provides essentially local information about surface orientation and does not require a large field of view. Thus, shape-from-texture-foreshortening mechanisms designed for general-purpose shape perception can be easily incorporated into a system for estimating planar surface orientation. Foreshortening information has, in fact, proven to be dominant in surface curvature studies in which it was pitted against other cues (density and its cousin, area, being most relevant for small field of view surfaces) [8,9,11,12,14].

9.4. Isotropy?

Given that the accumulating evidence suggests a strong role for foreshortening information in planar slant perception, we must consider the question of whether subjects can rely on an assumption of isotropy to interpret foreshortening information. The theoretical justification for weighing the foreshortening cue strongly relies on an assumption of surface texture isotropy. A number of previous studies tested

for an assumption of isotropy by compressing or stretching surface textures before projecting them and measuring resultant biases in subjects' estimates of surface orientation [10,15]. The results of these experiments are ambiguous, with predicted effects appearing in some conditions and not others and with large individual differences between subjects. Cumming et al. looked for an effect of isotropy by systematically compressing a surface texture before mapping it onto a curved surface and found that doing so greatly reduced the effectiveness of texture information relative to stereo in a curvature judgment task [11]. While consistent with the hypothesis that the visual system only relies on texture information when surface textures are isotropic, it should be noted that they compressed textures in a highly accidental way—the combination of viewing geometry and direction of compression led to texels being compressed in the direction of surface tilt. Other anisotropic textures may well be more informative about surface curvature.

The data from the experiments presented here; particularly those of experiments 2, 3 and 4 provides more unequivocal support that the subjects did use some form of isotropy constraint. The strongest evidence derives from the conditions in experiment 4 in which the subjects performed substantially better than the ideal observer which uses foreshortening information but does not rely on an assumption of isotropy.

In all of the present experiments, we used surface textures which were, in fact, isotropic. It may well be that the visual system relies on isotropy when it matches the stimulus information, but does not when the stimulus information is inconsistent with the constraint. This could explain the more equivocal results obtained when surface texture foreshortening was independently manipulated. More work needs to be done on this question. Of particular interest would be to look at how the relative weights of texture scaling and foreshortening information change when surface textures are made anisotropic. The ideal observer analyses, which show a strong diminution of foreshortening cue reliability with removal of the isotropy constraint, suggest that such a re-weighting would be computationally appropriate.

Acknowledgements

I am grateful to Zili Liu, who contributed to the initial formulation of the project, to Paul Schrater and Jack Nachmias for helpful discussions and to an anonymous reviewer who provided exceptionally detailed and helpful comments and suggestions on a first draft. This work was supported by NIH grant NEI-EY09383 to the author.

Appendix A

We used a 2D stochastic reaction-diffusion process to create samples of constrained random texel lattices, in order to avoid too-frequent overlap of texels. The model simulated a random process wherein particles migrate at random within a unit plane, subject to inhibitory fields surrounding each particle. For a surface containing N points, the motion of each point was described by the stochastic differential equation (Eq. (6))

$$\frac{\partial \vec{X}_i(t)}{\partial t} = \sqrt{T}\mathcal{N}(0,1) - \frac{1}{2}\nabla_i E(\{\vec{X}(t)\}) \tag{6}$$

where $\vec{X}_i$ is the position of point i, T is the temperature of the process, $\mathcal{N}(0,1)$ is a standard white Gaussian noise process and $E(\{\vec{X}\})$ is the total energy of the system at time, t, expressed as a function of the set of all point positions, $\{\vec{X}(t)\} = \{\vec{X}_0(t),\ \vec{X}_1(t),\ldots\vec{X}_N(t)\}$. $\nabla_i E(\{\vec{X}(t)\})$ is the gradient of E with respect to the position of point i. We modeled the energy as a sum of pair-wise interaction terms, $E = \Sigma_i \Sigma_j\ e_{i,j}$, where $e_{i,j}$ is given by (Eq. (7))

$$e_{i,j} = \begin{cases} (|\vec{X}_i - \vec{X}_j| - \delta)^2; & |\vec{X}_i - \vec{X}_j| < \delta \\ 0; & \text{Otherwise} \end{cases} \tag{7}$$

where δ is the spatial extent of the inhibitory field around each point. We set this to be the average spacing of a uniformly distributed set of points in the unit plane ($\delta = 1\sqrt{N}$). The noise component of the model forces points in random directions, while the inhibitory field around each point tends to force a regular spacing of points. The temperature parameter controls the degree of randomness in sample lattices—large values cause the noise term to swamp the inhibitory field, leading to highly irregular lattices, while small terms cause the inhibitory field to dominate, leading to very regular lattices.

Monte-Carlo simulations of the process described above (replacing the differential equation with a difference equation approximation) were used to generate samples of constrained random lattices. Each sample was generated by initializing the points using a 2D Poisson process, and then iterating through the stochastic difference equation until the energy of the system reached a statistical steady-state. At this point, the process was stopped and the point positions sampled to create a texel lattice. Since the system does not behave well with free boundary conditions, a toroidal geometry was assumed for the lattice, in which points near a boundary were treated as neighbors of points near the opposing boundary. The temperature for the simulations was set to $T = 0.01$.

Not only does the diffusion model support the creation of random texel lattices, it also supports the derivation of a prior probability model of texel positions, which was used to derive the ideal observer for the position cue.

References

[1] Gibson JJ. The perception of visual surfaces. Am J Psychol 1950;63:367–84.

[2] Braunstein M. Motion and texture as sources of slant information. J Exp Psychol 1968;78(2):247–253.

[3] Johnston E, Cumming B, Parker A. Integration of depth modules: stereopsis and texture. Vis Res 1993;33(5/6):813–26.

[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] Zimmerman GL, Legge GE, Cavanagh P. Pictorial depth cues: a new slant. J Opt Soc Am A 1995;12(1):17–26.

[6] Kraft AL, Winnick WA. The effect of pattern and texture gradient on slant and shape judgments. Perception Psychophy 1967;2:141–7.

[7] Braunstein M, Payne J. Perspective and form ratio as determinants of relative slant judgments. J Exp Psychol 1969;81(3):584–590.

[8] Cutting J, Millard R. Three gradients and the perception of flat and curved surfaces. J Exp Psychol Gen 1984;113(2):198–216.

[9] Todd J, Akerstrom R. Perception of three-dimensional form from patterns of optical texture. J Exp Psychol Hum Perception Performance 1987;13:242–55.

[10] Frisby JP, Buckley D. 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.

[11] Cumming BG, Johnston EB, Parker AJ. Effects of different texture cues on curved surfaces viewed stereoscopically. Vis Res 1993;33(5/6):827–38.

[12] Blake A, Bulthoff HH, Sheinberg A. Shape from texture: Ideal observers and human psychophysics. Vis Res 1993;33(12):1723–37.

[13] 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.

[14] 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, editors. Attention and Performance XVI: Information Integration in Perception and Communication. Hillsdale, NJ: L Erlbaum, 1996.

[15] Rosenholtz R, Malik J. Shape from texture: Isotropy or homogeneity (or both)? Vis Res 1997;37(16):2283–2293.

[16] Knill D. Surface orientation from texture: Ideal observers, generic observers and the information content of texture cues. Vis Res 1996;38(11):1655–1682.

[17] Witkin AP. Recovering surface shape and orientation from texture. Artif Intell 1981;17(1):17–45.

[18] Davis LS, Janos L, Dunn SM. Efficient recovery of shape from texture. IEEE PAMI 1983;5(5):485–92.

[19] Kanatani L, Chou T. Shape from texture: General principles. Artif Intell 1989;38:1–48.

[20] Blake A, Marinos C. Shape from texture: estimation, isotropy and moments. Artif Intell 1990;45:323–80.

[21] Kendall MK, Stuart A. The Advanced Theory of Statistics: Vol. 3. Design and Analysis, and Time Series. New York: Hafner, 1966.

[22] Rao C. Linear Statistical Inference and its Applications. New York: Wiley, 1973.

[23] Watson AB, Pelli DG. QUEST: A Bayesian adaptive psychometric method. Perception Psychophy 1983;33:113–20.

[24] 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.

[25] Fortune SJ. A sweepline algorithm for Voronoi diagrams. Algorithmica 1987;2(15):3–174.