Pergamon

PII: S0042-6989(96)00121-6

Vision Res., Vol. 37, No. 16, pp. 2283–2293, 1997

Printed in Great Britain
0042-6989/97 $17.00 + 0.00

Surface Orientation from Texture: Isotropy or Homogeneity (or Both)?

RUTH ROSENHOLTZ,*‡ JITENDRA MALIK†

Received 23 June 1995; in revised form 9 April 1996

We examine two models for human perception of shape from texture, based on two assumptions about the surface texture: isotropy and homogeneity. Observers made orientation judgments on planar textured surfaces. Surface textures were either isotropic or anisotropically stretched or compressed. If subjects used an isotropy assumption, they would make biased orientation estimates for the anisotropic textures. In some conditions some observers showed no bias for the anisotropic textures relative to the isotropic textures. In general, even when the observers showed bias, the biases were significantly less than those predicted if the observer used only deviation from isotropy as a cue. Observers appear to use both the deviation from isotropy and a texture gradient or affine texture distortion cue for shape from texture. © 1997 Elsevier Science Ltd.

Shape from texture Homogeneity Isotropy Texture

INTRODUCTION: TWO MODELS FOR SHAPE FROM TEXTURE

When we look at the image of a textured surface such as that shown in Fig. 1, we obtain a vivid percept of a plane slanted in depth. This pictorial cue has long been exploited by artists. However, its scientific study as a cue in visual perception started only with the seminal work of Gibson (1950).

Gibson coined the term *texture gradient* to describe the phenomenon that neighboring surface patches which have identical or sufficiently similar texture in the scene project in the retinal image to patches with different appearances due to differences in distance and surface orientation with respect to the viewer. Gibson used the term "gradient" to suggest measurement of some kind of change, though he did not have a mathematically precise way of characterizing that change. Subsequent research has resulted in the definition of a number of different texture gradients. We will illustrate them using Fig. 1 as a canonical example.

In this image, the *tilt* direction-defined as the direction in the image plane along which the distance to the viewed surface increases most rapidly-is vertical. Moving in the tilt direction, there is a change in the lengths of the major axes of the ellipses due to the fact that they are further away from the viewer. This is referred to as the *scaling* or *perspective gradient*. There is also a change in the aspect-ratio of the ellipses as we move in the tilt direction—the minor axes become smaller at a rate faster than the major axes. This is an instance of the *foreshortening* or *compression gradient*. Also, the areas of the ellipses decrease (the *area gradient*) and the density increases (the *density gradient*). The mathematical relationship of these different gradients to surface orientation and shape is well understood, both for planar surfaces (Stevens, 1981) and curved surfaces (Gårding, 1992). Closely related to the concept of texture gradients is the notion of *affine texture distortion*, which we have previously developed (Malik & Rosenholtz, 1994, 1997), in which the change in the texture is modeled locally as an affine

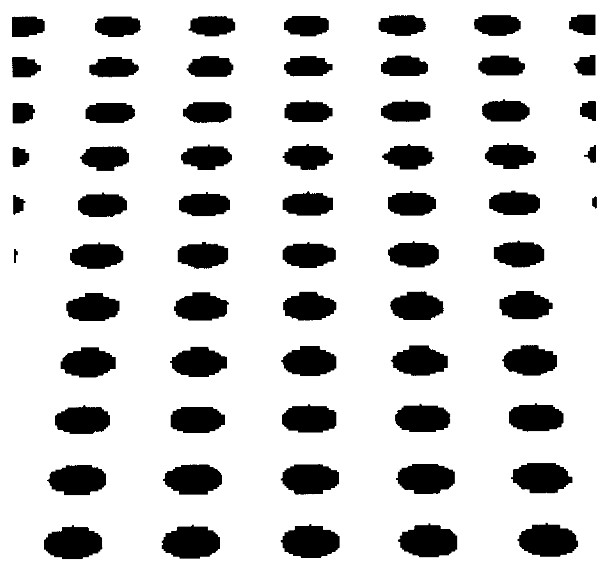

FIGURE 1. Planar surface textured with circular texture elements.

*Xerox PARC, 3333 Coyote Hill Road, Palo Alto, CA 94304, U.S.A.
†Department of Electrical Engineering and Computer Science, University of California at Berkeley, Berkeley, CA 94720, U.S.A.
‡To whom all correspondence should be addressed [*Email* rruth @parc.xerox.com].

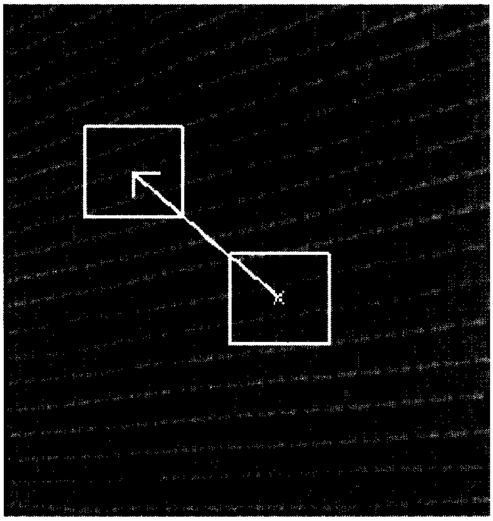

FIGURE 2. The texture distortion between two image patches can be modeled as an affine transformation: $[x', y']^T = A[x, y]^T + [\Delta x, \Delta y]^T$, where A is a 2×2 matrix. A depends on the local surface shape and orientation. A computational model of how this affine texture distortion can be used to recover the local surface geometry has been presented in (Malik & Rosenholtz, 1994, 1996).

transform (Fig. 2). This has the advantage that it subsumes the different texture gradients, and contains enough information to recover surface orientation and curvature locally.

In order for the measurements of texture gradients or affine texture distortion to specify surface orientation and shape, one must make some kind of *homogeneity* assumption about the surface texture. For instance, one could assume that the density of the texture was nearly constant on the surface, and use the way in which the density varies in the image to judge the shape and orientation of the surface (Marinos & Blake, 1990). Clearly, if the texture density on the scene surface itself varied in some contrived way, this cue would fail to lead to veridical perception. Similar remarks apply to the other texture gradients—we need to assume that that the density, area, foreshortening and/or other texture statistics are nearly constant, or *homogeneous*, on the surface. In the projection of this surface, these texture statistics will then vary only due to differences in projective distortion caused by changes in distance or orientation, enabling the visual system to use this variation to determine surface shape and orientation.

An entirely different class of models is based on a different assumption about the surface texture. If in Fig. 1, the visual system were to make the assumption that each ellipse was the projection of a circle lying on a slanted plane, it would be possible to locally infer the orientation of this plane without any need to measure texture gradients or distortion. Of course, assuming that texture elements are circles is not a general purpose shape-from-texture mechanism useful for natural scenes. Greater generality is obtained by the weaker assumption that the observer has knowledge about the statistics of the texture on the surface, and uses the deviation of the statistics of the image texture from those "known" statistics to determine the shape and orientation of the surface. Within this class of models the most common assumption is that the scene texture is *isotropic*, i.e. that it has no dominant direction or orientation bias (e.g. Witkin, 1981; Blake & Marinos, 1990). Under this assumption, the local foreshortening of the texture can be measured directly by measuring the deviation of the orientation distribution from isotropic. In Fig. 1, for instance, more of the orientation energy is distributed around the horizontal than the vertical direction. Image patches closer to the horizon are more slanted relative to the line of sight, and thus their orientation distribution deviates more from isotropic. For this particular texture, assuming an isotropic texture amounts to the same thing as assuming the texture elements are circles, but the isotropy assumption applies to a broader class of textures (Gårding, 1993).

Note the crucial difference between the use of this isotropy assumption and the use of a homogeneity assumption. With the homogenity assumption we need to *compare* two image patches and then use texture gradients or affine texture distortion as a cue to local surface orientation. Since such a model exploits the *change* between image patches, there is no need to assume that the orientation distribution is isotropic. The texture may be anisotropic; it is the *change* in the distribution from one image patch to another that is crucial, not the distribution itself.

In both models one assumes something about the surface texture (e.g. it is homogeneous, or it is isotropic), and then uses the deviation of the image texture from that assumption (e.g. texture gradients, or the deviation from isotropy) as a cue to the shape and orientation of the surface. In this paper we examine the two assumptions of isotropy and homogeneity in order to distinguish between these two models of shape from texture perception.

PREVIOUS WORK

Previous work has indicated that observers use some sort of texture gradient cue, at least for planar surfaces, and has suggested that observers might use deviation from isotropy as a cue, but has not clearly resolved the question of whether observers use one or both of these cues.

Cutting and Millard (1984) used a cue conflict paradigm to study which of the various texture gradients we use to perceive the "flatness" or "curvedness" of a textured surface. Subjects judged which of a pair of surfaces looked more like a flat slanted surface (or a curved surface, in the second experiment). They found that for planar surfaces 50–70% of the variance in the data was accounted for by the perspective gradient. In other words, observers had a strong impression of a slanted planar surface when it was indicated by a perspective gradient, but had a much weaker impression of a slanted surface when the perspective gradient was incorrect even though the deviation-from-isotropy cue always indicated a slanted surface. This implies that, at least for planar surfaces, observers do make use of a texture gradient type of cue. However, we cannot conclusively judge from their results whether or not observers also use a deviation-from-isotropy cue.

For curved surfaces Cutting and Millard found that the foreshortening gradient accounted for almost all of the variance in the data. However, since both the use of a foreshortening gradient cue and the use of a deviation-from-isotropy cue would predict this result, we cannot distinguish between the homogeneity and isotropy models from these results.

Todd and Akerstrom (1987), in their shape from texture experiments, concluded that subjects do not use a deviation-from-isotropy cue. They ran an experiment designed to compare "regular" and "irregular" textures. The regular texture consisted of square texture elements (or *texels*) of constant area, oriented randomly, which might overlap each other. In the irregular texture condition, the texels varied in area by up to a factor of three, with their lengths up to three times their widths.

They found no significant difference between the irregular and regular textures. They interpreted this result to mean that "observers do not perceive surfaces by assigning local depth values from optic element lengths or by assigning local orientation values from optic element compressions" (i.e. the isotropy model is incorrect), as argued by Stevens (1981, 1984) and Witkin (1981). However, this conclusion should be taken with a grain of salt. Because of the random orientations of the texels and the overlap between texels, the "regular" texture is already highly irregular, and overlapped square texels look a great deal like single, elongated texels. In fact, in their figure which compares surfaces with regular and irregular textures, the textures are indistinguishable in terms of regularity and anisotropy.

In other experiments, Todd and Akerstrom demonstrated that observers perceived a greater amount of depth when the texture elements were elongated perpendicular to the tilt direction, even when the foreshortening of the texels was held constant over the image. They interpreted this result as evidence for their model of shape from texture, in which early stages emphasize oriented texels with similar orientations. However, it could perhaps also be explained by the use of an isotropy assumption, as noted by Cumming *et al.* (1993).

Both the conclusions of Cutting and Millard and those of Todd and Akerstrom, with regard to what texture gradients observers use to judge shape from texture, should be drawn with caution because of the cue conflict nature of the experiments. In a cue conflict situation, if an observer does not see shape from texture, it could be because the observer does not use the cue which correctly indicates the shape, or it could be that the conflicting information from the other, "incorrect" cues may destroy the percept. This is less of a problem in drawing conclusions from their results about the use of a texture gradient type of cue vs an deviation-from-isotropy cue. However, the conclusions are still questionable, because the method which an observer uses to determine shape for an image with conflicting cues, *that would be unlikely to exist in normal everyday life*, may differ greatly from the method typically used.

Cumming *et al.* (1993), asked subjects to judge the depth of cylinders textured with both isotropic and anisotropic textures. They show that observers perceive less depth for their more elongated, anisotropic, ellipse textures than for either isotropic circular textures or isotropic texture formed by randomly orienting elongated ellipses. From this, they conclude, "that human shape-from-texture works on the assumption that surfaces are covered with approximately isotropic textures." However there are alternative explanations of the poorer performance on anisotropic textures than on isotropic textures. The anisotropic textures may simply provide less information than the isotropic textures; as they point out, their more anisotropic textures have more variance in their aspect ratios. Furthermore, for highly anisotropic textures (their textures have aspect ratios as high as 3.0) one must detect much smaller changes in element foreshortening, which could explain the poorer performance relative to both kinds of isotropic textures.

Blake *et al.* (1993) used an ideal observer model for shape from texture to compare a model in which observers use an isotropy assumption with one in which observers assume the texture has constant density on the surface and use the density gradient to perform shape from texture. For their textures, they used line segments which were randomly oriented with a uniform distribution over 180°. The line segments varied in length up to a factor of 2. Observers judged the shape of textured cylindrical surfaces. Blake *et al.* (1993), then determined the information available from the density gradient, from the deviation from isotropy, and from both combined, for determining the shape from texture. The information content is in the form of predicted variance in the shape estimates. They compared this predicted variance in the shape estimates to their experimental results. Using this methodology, Blake *et al.* (1993), showed that the visual system must make use of cues other than the density gradient, because observers performed better at shape from texture than they could using the density gradient alone. However, while they showed that observers must use *more* than the density gradient, they did not actually show that observers use the density gradient at all. Furthermore, it is not completely clear that the additional cue was the deviation from isotropy. The change in compression of the texture, rather than an assumption of textural isotropy, might have provided the additional information, as might the change in average lengths of the line segments.

In conclusion, it is fair to say that there has not yet been a decisive answer to the question: Do humans use homogeneity or isotropy (or both) in order to infer surface orientation from texture?

VORONOI POLYGON TEXTURES

Our core idea is simple: if we ask subjects to make orientation judgments on surfaces textured with anisotropic textures, subjects should, if they use the deviation of the image texture from isotropy, give biased estimates of the surface orientation relative to their estimates of surface orientation for isotropically textured surfaces. If subjects use only a texture gradient type of cue there should be no bias in their estimates for anisotropically textured surfaces (at least for a reasonable range of anisotropy).

We need to define a suitable set of stimuli for which one can easily control the amount of anisotropy. Furthermore, textures such as that shown in Fig. 1, with regular placement of the texture elements, can lead to global orientation cues which are not modeled by local models of shape from texture such as either the isotropy or homogeneity models discussed here. Finally, in addition to wanting irregular placement of the texels, we wanted the "texels" themselves to be fairly irregular, so that no particular feature of a "texel" would "point" in the tilt direction, as would be the case for textures composed of familiar shapes such as circles, ellipses, and rectangles.

An analogy with random dot stereograms and kinematograms is appropriate. Just as RDSs and RDKs have been designed specifically to try and isolate low level mechanisms for stereopsis and motion, we would like to devise texture stimuli that avoid conflicts from cues due to familiar forms.

For our experiments, we introduce a novel class of stimuli, *Voronoi Polygon Textures* with a number of advantages for the psychophysical study of shape-from-texture. Figure 3 shows a typical Voronoi Polygon texture, mapped onto a frontoparallel plane. These textures are based on the concept of a *Voronoi diagram* (Aurenhammer, 1991) of a set of points on a plane. Given a set of points, or *sites*, on a plane, a Voronoi diagram divides the plane into a set of *Voronoi polygons*, one polygon per site, such that all points in a polygon are closer to the site corresponding to that polygon than to any other site. To create our textures, we first compute the Voronoi diagram for a given set of points using the algorithm of Fortune (1987). This gives us a set of *Voronoi polygons*. Scaled-down versions (in our experiments, by a factor of 0.8) of the Voronoi polygons are the texels which make up our texture.

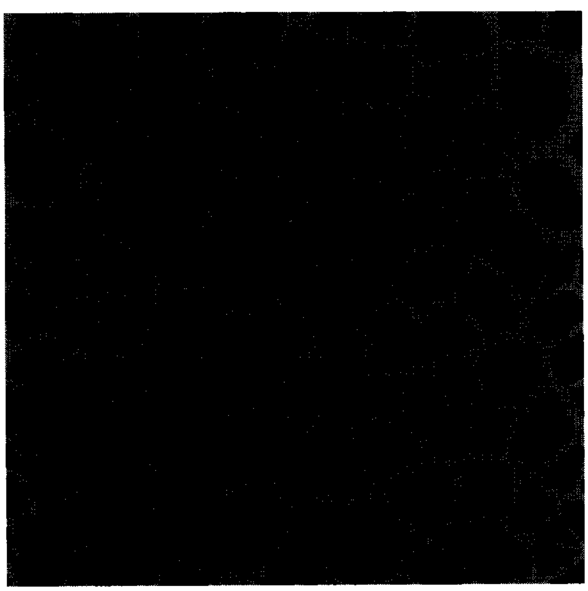

FIGURE 3. Example isotropic Voronoi polygon texture.

Voronoi polygons allow us to create natural-looking irregular textures, since many natural textures resemble Voronoi diagrams (Aurenhammer, 1991); for instance, whenever one has a number of items, such as cells, which all start growing at roughly the same time, and grow at the same rate until they run into each other.

A number of different parameters control the appearance of a Voronoi polygon texture. First, we can control the spatial placement of the sites. Typically, the sites will be generated as a realization of a random spatial point process (Stoyan & Stoyan, 1994). By varying the point process which generates the location of the texels we can test a full range of textures from extremely irregular textures to fairly regular textures.

The canonical example of a spatial point process is the Poisson process, which corresponds to complete spatial randomness. Formally, a Poisson process is characterized by the property that for any disjoint regions, $B_1, \ldots, B_k$, the numbers of points in these regions, $N(B_1), \ldots, N(B_k)$, are stochastically independent. $N(B)$ is a Poisson random variable with expected value λ*Area(B), where the parameter λ denotes the intensity or the mean point density. We can simulate a realization of a Poisson process in a region by dropping points at random in the region, where each new point can be anywhere in the region with equal probability.

A number of different spatial point processes have been defined in the literature (Stoyan & Stoyan, 1994) to model various spatial phenomena. We chose the class of point processes defined by so-called *hard-core* models (Fig. 4). The distinction between these models and Poisson processes is that in a realization of a hard-core model no two points may lie less than a distance $2R$ apart. The parameter R defines an inhibition zone around a point. One example of such a process would be dropping

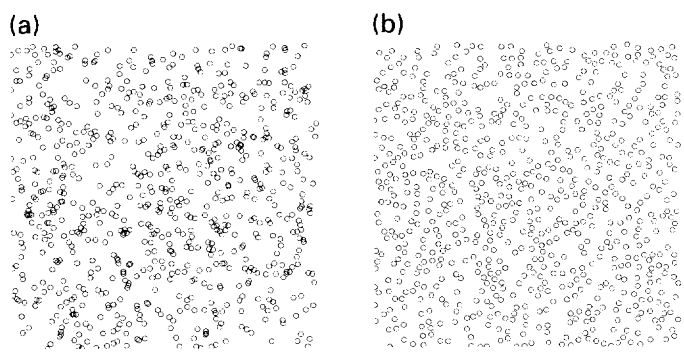

FIGURE 4. Two realizations of hard-core models with identical λ. The inhibition radii are different: 0.002 (a) and 0.01 (b), where we depict a unit area. Note the regularity of the second texture compared to the first.

marbles onto a planar surface: the marbles can land anywhere with equal probability, but not on top of one another. This is a more realistic model than a Poisson process for many natural textures; plants, for instance, do tend not to grow on top of each other. As the density increases, the realizations of these processes start looking more and more regular and assume quite a periodic appearance. Varying the inhibition radius, R, for a given density offers a technique for generating textures varying on a continuum from regular to irregular, as seen in Fig. 4. For low densities relative to the "texel size" (inhibition radius) the process is extremely irregular, and approaches a Poisson process. For higher densities relative to the inhibition radius, the inhibition requirement makes the textures approach a regular appearance.

Given a set of random sites, e.g. those generated by a hard-core model, that define the Voronoi diagram, the diagram will itself have fairly randomly shaped and positioned Voronoi polygons, and thus will not give global cues to surface orientation. Given isotropically distributed sites, we get an isotropic texture. Our Voronoi polygon textures give a strong impression of slant, in spite of their irregularity.

In addition to varying the inhibition radius of the hard-core model, a number of other control parameters are available for psychophysical studies. One could scale the Voronoi polygons and independently vary the size of the texels. One could replace the Voronoi polygons with texels with the same area as the polygons but different shape, e.g. to get a texture with uniformly-shaped elements.

DO OBSERVERS USE HOMOGENEITY OR ISOTROPY (OR BOTH)?

Overview

We asked subjects to indicate the perceived orientation of a textured planar surface for a number of different orientations and for both isotropic and anisotropic textures. If subjects make use of an isotropy assumption to find the orientation of a surface, we expect to see bias in their orientation estimates for the anisotropic textures.

We parametrize the surface orientation using two parameters: slant and tilt. The *tilt* is the direction in which the distance to the surface changes most rapidly. This is also the direction of the projection of the surface normal in the image plane. The *slant* is the amount by which the surface orientation differs from frontoparallel; it is the angle between the line of sight and the surface normal. Stevens (1983) has argued for the advantages of this representation of surface orientation.

We wish to answer the following questions:

- Do subjects overestimate slant if the texture is compressed in the tilt direction, as predicted by an isotropy assumption?
- Do subjects underestimate slant if the texture is stretched in the tilt direction?
- If we compress the texture in a direction not aligned with the tilt direction, do subjects show biases in slant and tilt as predicted by an isotropy assumption?

If subjects do show biases in their orientation estimates, we would like to determine whether their biases are as large as would be predicted if they used only the deviation-from-isotropy cue, or if subjects seem to combine this cue with other shape from texture cues such as the texture gradient or texture distortion cues.

Note that if people do use an isotropy assumption, then we have a cue conflict in our stimuli. However, this cue conflict is not as serious as that in previous studies because anisotropic textures exist in the real world, and the human visual system should be able to deal with them in the same way as it would with textures in the natural environment.

In addition, note that we make no assumptions about just how subjects might *measure* the anisotropy of the image texture, i.e. whether this task might be done as a low-level measurement of orientation content, or as a higher level process in which polygons are first extracted, and then their mean aspect ratio measured. Similarly, we make no assumptions about whether subjects use one of the several texture gradients (Gårding, 1992) or the affine texture distortion (Malik & Rosenholtz, 1994, 1997), if they make use of a homogeneity assumption. These are interesting questions, but not the ones we address here.

Experimental design

Subjects viewed images of perspectively projected, slanted, textured planes displayed on a Silicon Graphics Indigo, through a circular window 21.2 cm (36 deg*) in diameter cut in black poster board. The window kept subjects from seeing the horizon of the slanted planes. Subjects sat at a distance of 32.6 cm, with their chin in a chin rest, resulting in 20 pixels/deg. The viewing distance was such that the projection onto the retina agreed with the projection used to generate the image.

Figure 3 shows a typical texture, mapped onto a frontoparallel plane. We generated approximately 0.07 sites/deg^2, according to a hard-core model with inhibition radius of 0.6 deg. This created a fairly regular texture.

The subjects viewed the stimuli monocularly. They indicated the perceived orientation of the plane by adjusting a gauge figure in the center of the image. The gauge, modeled after that used by Koenderink *et al.* (1992), had a diameter which measured 3.5 deg and consisted of a red circular disk with a needle, perpendicular to the disk, passing through the center of the disk.† The needle was half green and half blue. The initial position of the gauge figure was chosen randomly on each

*To avoid confusion, "deg" is used to denote stimulus size and for "°" slant and tilt angles.

†Our gauge is slightly larger than that used by Koenderink *et al.*, which had a diameter which measured 2.5 deg.

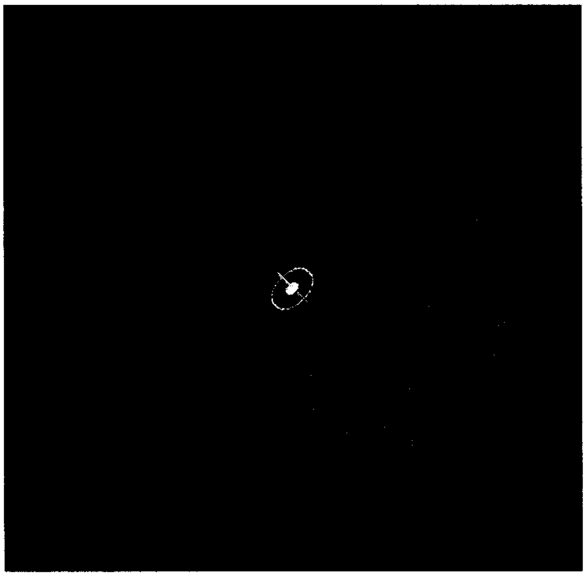

FIGURE 5. Sample image for Experiments 1–3, showing gauge for indicating orientation. The actual gauge is red, with a green and blue needle.

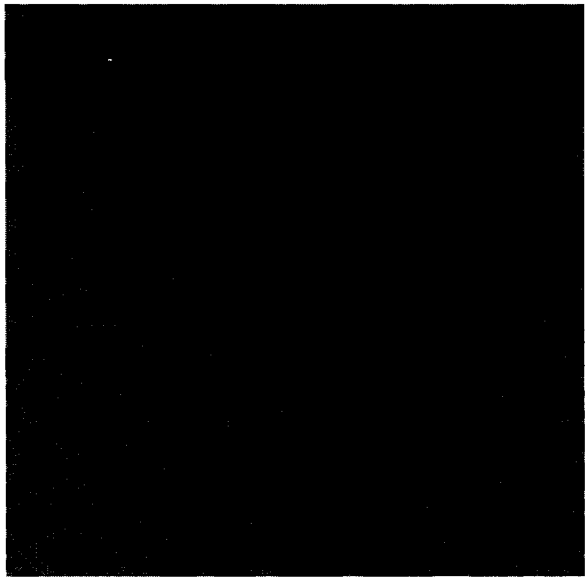

FIGURE 6. Example "compressed" Voronoi polygon texture.

trial. The subjects' task was to align the gauge so that it looked like the circular disk laid on the textured surface, with the green portion of the needle pointing in the direction of the surface normal on the side of the surface towards the observer. Subjects used the computer mouse to adjust the orientation of the gauge figure. Figure 5 shows a typical textured surface, with a grayscale version of the gauge.

We put no limit on how long subjects could take to make their orientation judgments. Subjects typically made 450 orientation judgments in ≈ 30 min.

Subjects participated in a training phase, in which they made judgments on the orientation of 50 surfaces. For the training phase, the surfaces were textured with a texture consisting of randomly placed rectangles of constant size. The orientation of the surface was chosen randomly from slants between 0° and 50°, and tilts between 0° and $-180°$. As in the actual experiments, subjects adjusted the gauge until they perceived it to be at the correct orientation, and then pressed the space bar to record that orientation. During the training phase, the subjects were then shown the correct orientation of the gauge, to provide feedback. Subjects received no feedback during the actual experiment.

We ran the experiment on three naive subjects, with normal or corrected to normal vision.

Stimuli

We generated five different textures for each of four classes of texture:

1. Isotropic texture.
2. Anisotropic texture that has been "compressed" in the tilt direction.
3. Anisotropic texture that has been "stretched" in the tilt direction.
4. Anisotropic texture that is compressed at an angle of 45° from the tilt direction, so that the anisotropy is not aligned with the tilt. (The texture is compressed by the same amount as Texture 2.) We call such textures *non tilt-aligned.*

The experimental sessions were blocked, with only one class of texture per session. The order in which the subjects saw the four classes of texture was randomly chosen for each subject.

In all cases we generated the Voronoi polygon textures described earlier, and then, for the anisotropic cases, compressed or stretched them prior to applying them to the planar surface. Figures 6 and 7 show typical compressed and stretched Voronoi polygon textures, such as those used to generate the tilt-aligned anisotropic textures. Figure 8 shows a typical non tilt-aligned

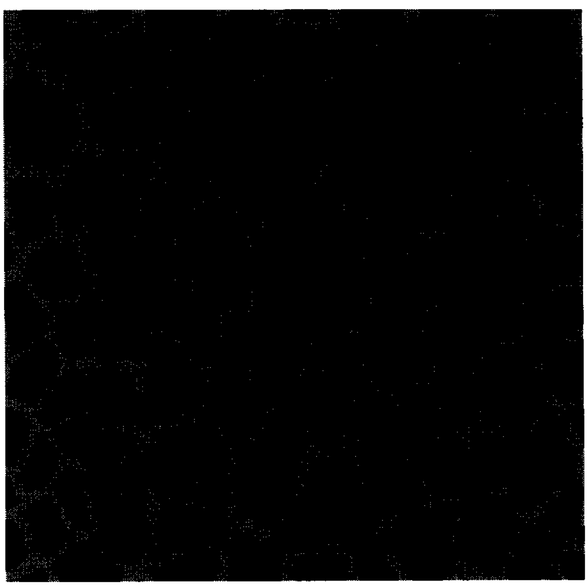

FIGURE 7. Example "stretched" Voronoi polygon texture.

FIGURE 8. Example texture, compressed at 45° from the tilt direction. Tilt = −90°.

anisotropic texture, mapped onto a plane with a tilt of −90. (On a frontoparallel plane, this texture looks the same as that in Fig. 6.)

If subjects make use of an isotropy assumption to find the orientation of the surface, we expect to see bias in their estimates of slant for the tilt-aligned anisotropic textures, and biases in tilt and slant for the nonaligned anisotropic texture, relative to their orientation judgments for the isotropically-textured surfaces. In particular, the amount of compression or stretch was such that if the observers used solely the anisotropy of the projected texture in the center of the image (near the gauge) to judge the surface orientation, then:

- For the texture compressed in the tilt direction, we would expect the observers to overestimate slant by 15°.
- For the texture stretched in the tilt direction, we would expect the observers to underestimate slant by 15°.
- For the texture compressed 45° from the tilt, we would expect the slant and tilt biases shown in Fig. 9. Roughly speaking, we would expect an increase in slant estimates by about 7°, and a tilt bias of between −20° and −26°.

This gives us our textures a maximum aspect ratio of 1.52 (prior to projection). This is comparable to the smallest amount of anisotropy in the anisotropic textures of Cumming *et al.* (1993).

Because of the possibility that subjects would underestimate slant by as much as 15° for the case of the stretched texture, we used only surfaces with slants of 15° or greater in the experiment. For the isotropic texture and the tilt-aligned anisotropic texture, subjects made orientation judgments on surfaces with slants of 15°, 20°, 30°, 40°, and 50°.

For the non tilt-aligned anisotropic textures we were predominantly interested in whether or not we would see the tilt biases predicted by an isotropy assumption. As we expected high variance in the tilt estimates for low values of slant (for a slant of 0° the tilt is actually undefined), we did not use slants <30° because for lower values of slant we expected that it would be difficult to say anything statistically significant about the tilt bias. Subjects therefore made orientation judgments on slants between 30° and 50°, in steps of 5, for this type of texture.

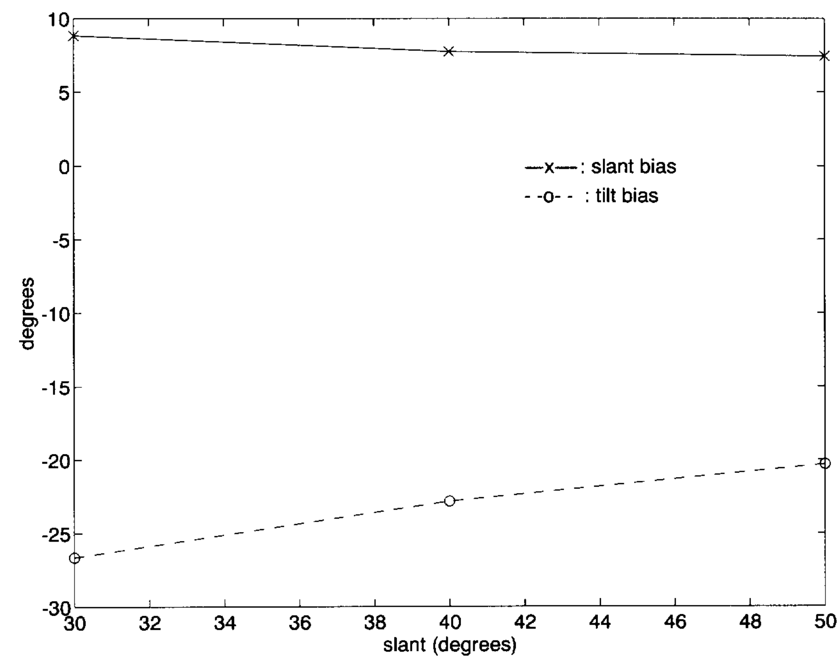

FIGURE 9. Expected bias for slant and tilt under the assumption of isotropy, for texture elongated 45° off from the tilt direction.

In all cases the tilts ranged from 0° to −170° in steps of 10°. The slant and tilt combinations occurred in random order, with the condition that no slant or tilt appear twice in a row. The order of the slant, tilt, and texture combinations were randomly generated in advance and stored in a file read in at the beginning of the experiment. We generated five different textures for each class of texture. For each class of textures, the subjects saw each of the five textures at each of the possible orientations, and thus subjects made 450 orientation judgments for each of the four types of texture.

RESULTS AND ANALYSIS

Preliminary analysis suggested that subjects' results did not depend on the tilt direction, over the range of tilts used. Therefore in the results that follow we average over different values of tilt.

Anisotropies aligned with tilt

Figure 10 plots the results for the three subjects, for the tilt-aligned anisotropic textures.

We had predicted that if subjects used only the anisotropy of the texels near the gauge figure to judge shape from texture, then we would expect a 15° increase in slant estimates for the compressed textures and a 15° decrease for the stretched textures, over the estimates for the isotropic textures.

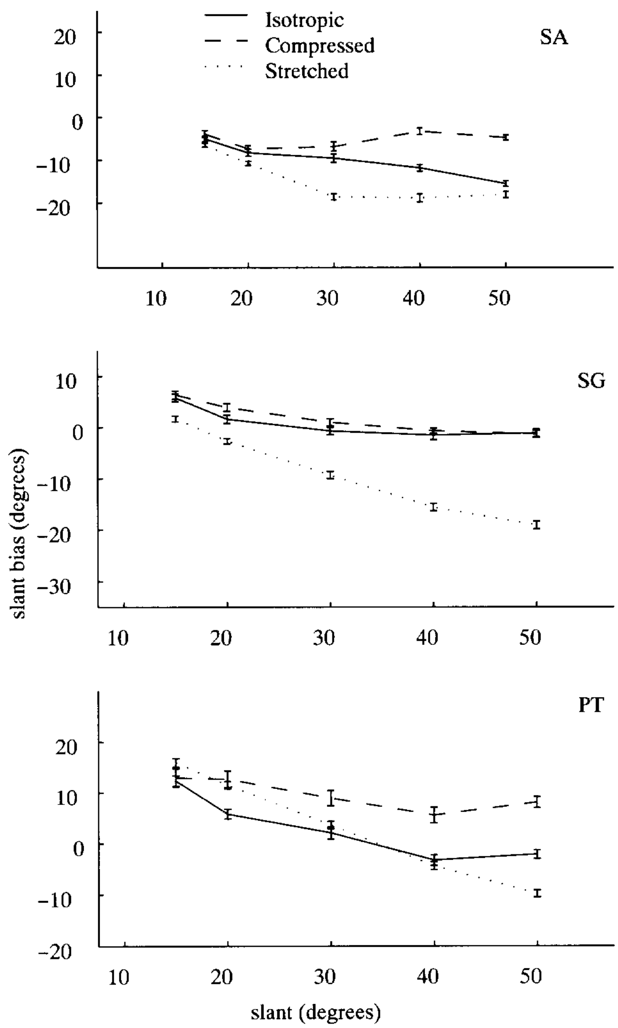

FIGURE 10. Comparing slant bias for isotropic and anisotropic textures.

Subject SA showed both an increase in estimated slant for the compressed texture [$t(178) = 1.91, 7.77, 11.98$, for slants of 30°, 40°, and 50°, respectively; $p < 0.05$ or better], and a decrease in estimated slant for the stretched texture [$t(178) = 1.76, 2.67, 7.52, 5.79, 2.57$, for slants of 15°, 20°, 30°, 40°, and 50°; $p < 0.05$ or better], though both the increase or decrease were by significantly less than the 15° predicted if the subject used only the deviation-from-isotropy cue [$t(178) = 8.57, 5.91, 4.73$, for the compressed texture; $p < 0.05$ or better. $t(178) = 17.0, 13.3, 4.99, 6.55, 12.5$, for the stretched texture; $p < 0.01$].

Subject SG generally performed the same for the compressed texture as for the isotropic texture, showing a small significant increase only for a slant of 20°. This subject showed significant decrease in slant estimates for the stretched texture [$t(178) = 4.30, 4.37, 8.31, 11.96, 16.67$, for slants of 15°, 20°, 30°, 40°, and 50°, respectively; $p < 0.01$]. For slants of 15°, 20°, and 30°, this decrease in slant estimates was by less than the 15° predicted if the deviation from isotropy were the only cue [$t(178) = 11.37, 10.90, 6.18$, respectively; $p < 0.01$]. For a slant of 50° the subject actually showed a decrease in estimated slant by *more* than the predicted 15° [$t(178) = 2.78$, $p < 0.01$], which may suggest that the cue conflict caused the subject to have difficulty perceiving the surface at that slant.

Subject PT, on the other hand, showed an increase in estimated slant for the compressed texture [$t(178) = 3.68, 3.59, 4.74, 7.46$, for slants of 20°, 30°, 40°, and 50°; $p < 0.01$], but no consistent decrease in estimated slant for the stretched texture. Again, the increase in estimated slant for compressed texture was significantly <15° [$t(178) = 4.33, 4.24, 3.29, 3.47$, $p < 0.01$]. For stretched textures, the significant *increase* in estimated slant for slants of 15° and 20° [$t(178) = 2.69, 4.74$; $p < 0.01$], and significant decrease in estimated slant for a slant of 50° [$t(178) = 7.11$, $p < 0.01$], may suggest that this subject could perceive the anisotropy of the texture for lower slants, and so adopted a different strategy for estimating surface orientation. The stretched texture should be easier to perceive when the surface is closer to frontoparallel, since some of the texels may project to look isotropic, or even anisotropically stretched.

The significant increase (decrease) in slant estimates for compressed (stretched) texture for subjects SA and PT (SA and SG), but by less than the 15° predicted by an isotropy model suggests that subjects use the deviation-from-isotropy cue, but combine it with some cue like texture gradients or affine texture distortion. The variability among subjects is similar to that typically found in subject response to a cue conflict situation by other psychophysicists.

We were concerned about a possible confounding factor in our data analysis that could bring into question the significance of the difference between the subjects' biases for anisotropic textures and the biases predicted by the isotropy model. Subjects indicated a compressed range of slant values—they tended to overestimate slant for low values of slant, and underestimate slant for higher values of slant. It is unclear whether this is due to a compressed range of perceived slant, or to a compressed range of indicated slant. The phenomenon of slant "regression to the fronto-parallel plane" has been noted before, so this is not a novel observation.

The problem in our data analysis arises because it is possible that subjects could have shown less bias than predicted by the use of an isotropy cue because of this compressed range of slant. We dealt with this as follows: we made new predictions of the isotropy model which take into account the compressed range of indicated slants by using the orientation estimates for the isotropic texture as a look-up table which describes the compression of the slants. Take the case of the anisotropically compressed texture: At a slant of 15°, the isotropy model predicts that the observer will see a slant of 30°. So we look up the mean estimated slant, for the isotropic texture, for a slant of 30°. This is the predicted slant estimate. From this we can compute the predicted slant bias, and compare this with the actual bias, as before. (For, a slant of 20°, e.g. we interpolate between the mean estimated slants of 30° and 40° to get an estimate of the estimated slant for a slant of 35°.)

We show these new predictions along with the data in Fig. 11.

Using this method we cannot get new predictions in all cases, given our data. For example, for a slant of 50°, we would have to look up the estimated slant for a slant of 65°, which would require data extrapolation. Of the 17 cases in which, according to our original predictions, subjects made orientation estimates which were significantly different from both the homogeneity-alone and isotropy-alone prediction, there are eight cases for which we cannot make new predictions from our data. Of the remaining nine cases for which we can make new predictions, in three of the cases the new predictions of the isotropy model no longer allow us to say that the estimates are significantly different from the isotropy-alone prediction. In the last six cases, the slant estimates still differ significantly from the new predictions of the isotropy model. [Subject SA, compressed texture,

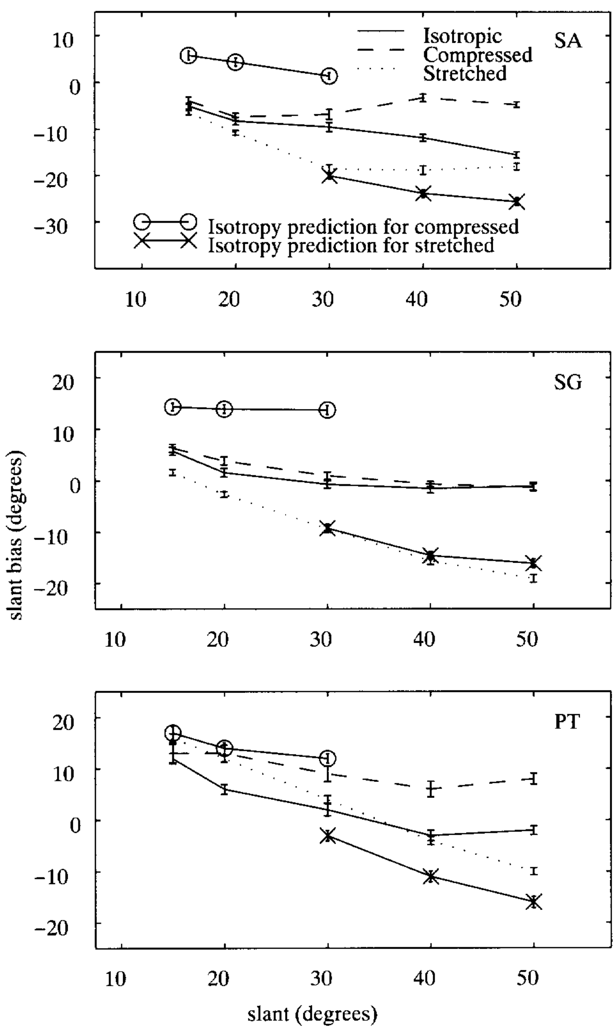

FIGURE 11. New predictions for slant bias.

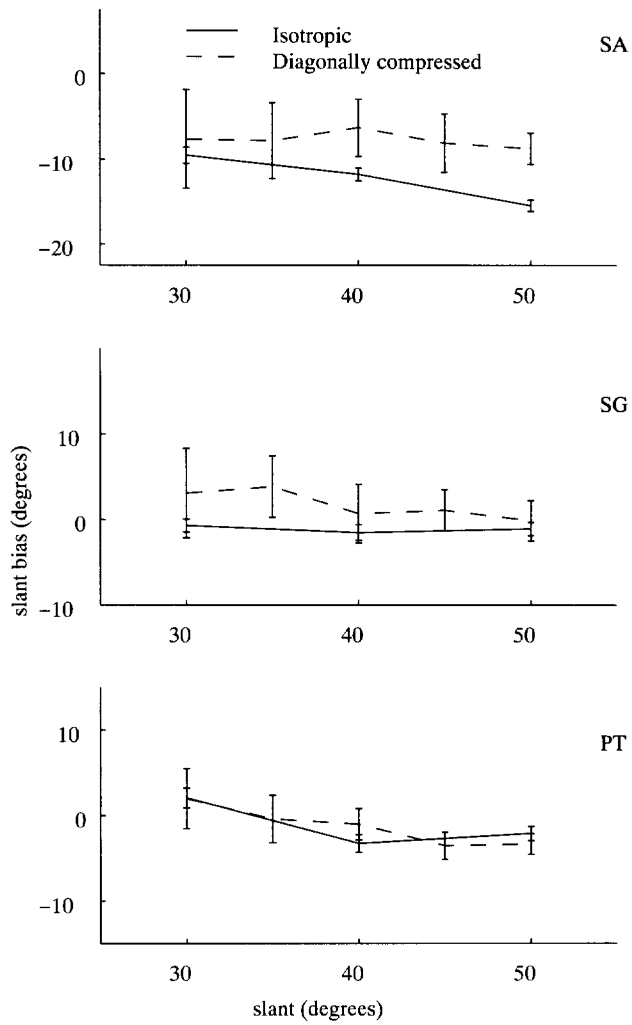

FIGURE 12. Comparing slant biases for isotropic and nontilt-aligned anisotropic textures.

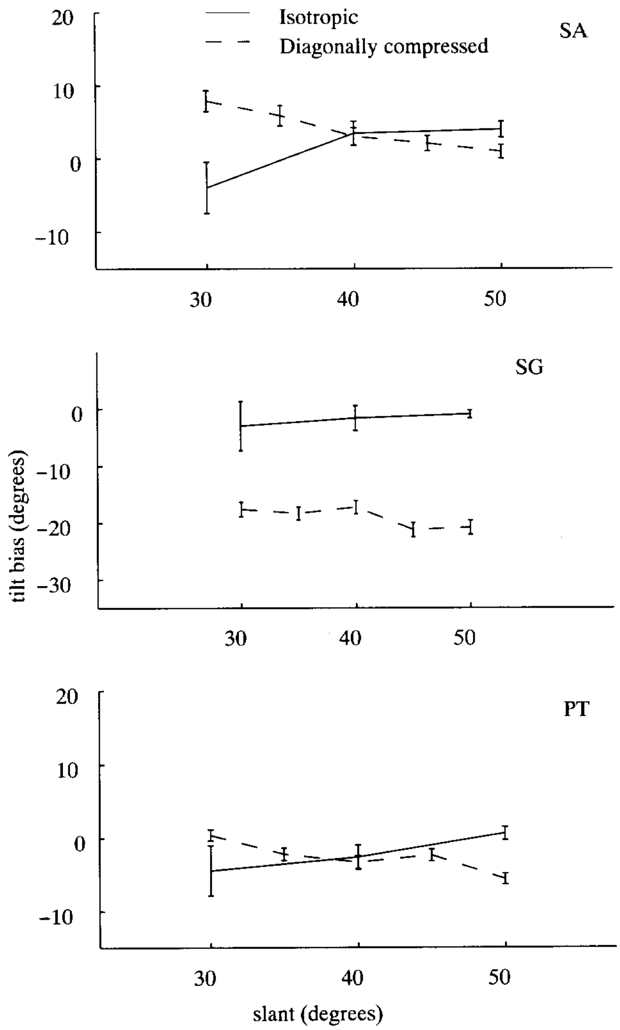

FIGURE 13. Comparing tilt biases for isotropic and nontilt-aligned anisotropic textures.

$t(178) = 4.19$ for a slant of 30, $p < 0.05$ or better. Subject SA, stretched texture, $t(178) = 3.95$, 6.73 for slants of 40° and 50°, $p < 0.05$ or better. Subject SG, compressed texture, $t(178) = 8.88$ for a slant of 20°, $p < 0.05$ or better. Subject PT, compressed texture, $t(178) = 1.89$ for a slant of 30°, $p < 0.05$. Subject PT, stretched texture, $t(178) = 4.46$ for slant of 50°, $p < 0.05$ or better.]

Anisotropies not aligned with tilt

For the nontilt-aligned anisotropic textures, an isotropy assumption predicted both an overestimate in slant by roughly 7°, and a bias in tilt between −20° and −26°. Figure 12 compares the results of slant judgments for the nontilt-aligned anisotropic textures to the results for isotropic textures. Figure 13 shows the results of tilt judgments for the anisotropic and isotropic textures.

Again, we see variability among subjects. We see significant increases in slant estimates only for subject SA, and only for a slant of 50° [$t(178) = 3.45$, $p < 0.01$]. This increase was not statistically different from the increase predicted by the isotropy model.

We see the predicted decrease in tilt estimates only for subject SG [$t(178) = 9.67$, 10.50, 13.37, for slants of 30°, 40°, and 50° respectively; $p < 0.01$]. For slants of 30° and 40°, the difference between this decrease and that predicted by the isotropy model was statistically significant [$t(178) = 7.94$, 4.84, $p < 0.01$]. Once again, this suggests that some subjects do use the deviation from isotropy as a cue, in combination with a cue based on texture gradients or affine texture distortion.

DISCUSSION

Our results suggest that subjects make use of *both* deviation from isotropy and texture gradient/distortion as cues for shape recovery. We can conclusively *dismiss* both of the following two hypotheses:

1. Subjects use only texture gradients/distortion.
2. Subjects use only the deviation from isotropy.

The data from any of the three subjects could be used to reject these two extreme hypotheses. While both cues are used, there is a good deal of variability among subjects on their cue combination strategies. Some subjects, in some conditions, do not seem to use an isotropy assumption, while under different conditions, or with different subjects, we do see a bias in orientation estimates for the anisotropic textures, suggesting that in these conditions the subjects do make use of a deviation-from-isotropy cue. It appears, furthermore, that even in the conditions in which subjects do make use of an isotropy assumption, their biases are typically smaller than those predicted by the isotropy model alone. This suggests that subjects in these conditions combine the deviation-from-isotropy cue with a texture gradient or distortion type of cue.

We conclude from our results that one needs to incorporate both the deviation-from-isotropy cue and a texture gradient/affine texture distortion cue in any complete shape from texture model. Initially, this might seem to be a surprising strategy for the visual system to follow: the homogeneity assumption is usually well-justified, whereas isotropy is often violated in particular scenes, even though it may true for an ensemble of scenes. Consider a field of largely vertical blades of grass, for instance. One possible speculation might be that the visual system makes a bias-variance tradeoff. The use of an isotropy assumption may reduce the variance of the estimate at the expense of introducing biases.

For a complete shape from texture model we need to understand how to combine these two cues. This necessitates further studies analogous to those done by Young *et al.* (1993) in the combination of texture and motion cues. One might expect the combination rule for an isotropy- and homogeneity-based shape from texture cues to be complex; the strength of each cue is affected by the field of view, the shape of the surface, the degree of anisotropy, the regularity of the texture, and so on. For planar surfaces receding in depth, and for the fairly large field of view we used, the stimuli had fairly large perspective effects. Since increased perspective aids the use of a texture gradient sort of cue, one expects that such a cue might be more informative and assigned more weight under such conditions than it might be assigned given a smaller field of view. How do subjects weight the two cues for curved surfaces? Furthermore, given an obviously anisotropic texture, do subjects cease to use the isotropy cue? Finally, we might expect that for more regular textures the texture distortion cue might dominate over the isotropy cue, in which case we would expect to see less effect from using anisotropic textures. All of these issues remain topics for future research.

REFERENCES

Aurenhammer, F. (1991). Voronoi diagrams—a survey of a fundamental geometric data structure. *ACM Computing Surveys, 23*, 345–405.

Blake, A., Bulthoff, H. & Sheinberg, D. (1993). Shape from texture: Ideal observers and human psychophysics. *Vision Research, 33*, 1723–1737.

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

Cumming, B. G., Johnston, E. B. & Parker, A. J. (1993). Effects of different texture cues on curved surfaces viewed stereoscopically. *Vision Research, 33*, 827–838.

Cutting, J. & Millard, R. (1984). Three gradients and the perception of flat and curved surfaces. *Journal of Experimental Psychology: General, 113*, 198–216.

Fortune, S. J. (1987). A sweepline algorithm for Voronoi diagrams. *Algorithmica, 2*, 153–174. This code is available via anonymous ftp from netlib.att.com, in the netlib/voronoi directory.

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

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

Gibson, J. (1950). *The perception of the visual world.* Boston: Houghton Mifflin.

Koenderink, J. J., Van Doorn, A. J. & Kappers, A. M. (1992). Surface perception in pictures. *Perception and Psychophysics, 52*, 487–496.

Malik, J. & Rosenholtz, R. (1994). Recovering surface curvature and orientation from texture distortion: a least squares algorithm and sensitivity analysis. *Proceedings of ECCV*, Stockholm, published in *Lecture Notes in Computer Science, 801*, 353–364.

Malik, J. & Rosenholtz, R. (1997). Computing local surface orientation and shape from texture for curved surfaces. *International Journal of Computer Vision, 23(2)*.

Marinos, C. & Blake, A. (1990). Shape from texture: The homogeneity hypothesis. *Proceedings of ICCV*, Osaka, Japan, pp. 350–353.

Stevens, K. A. (1981). The information content of texture gradients. *Biological Cybernetics, 42*, 95–105.

Stevens, K. A. (1983). Slant-tilt: The visual encoding of surface orientation. *Biological Cybernetics, 46*, 183–195.

Stevens, K. A. (1984). On gradients and texture "gradients". *Journal of Experimental Psychology: General, 113*, 217–220.

Stoyan, D. & Stoyan, H. (1994). *Fractals, random shapes and point fields: Methods of geometrical statistics*. Chichester: Wiley.

Todd, J. T. & Akerstrom, R. A. (1987). Perception of three-dimensional form from patterns of optical texture. *Journal of Experimental Psychology: Human Perception and Performance, 13*, 242–255.

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

Young, M. J., Landy, M. S. & Maloney, L. T. (1993). A perturbation analysis of depth perception from combinations of texture and motion cues. *Vision Research, 33*, 2685–2696.

Acknowledgements—The work described in this paper was supported in part by NSF PYI Grant IRI-8957274, and JSEP Contract F49620-93-C-0014. We would like to thank Karen DeValois and Marty Banks for their helpful comments. We thank Andrea Mennucci at Caltech for his computer program for displaying the gauge figure, and Dionne Boyce for help with part of the computer implementation.