Starry night: a texture devoid of depth cues

Xenophon Zabulis

*Institute for Research in Cognitive Science, University of Pennsylvania, 3401 Walnut Street,
Philadelphia, Pennsylvania 19104-6228*

Benjamin T. Backus

Department of Psychology, University of Pennsylvania, 3401 Walnut Street, Philadelphia, Pennsylvania 19104-6228

From a modern Bayesian point of view, the classic Julesz random-dot stereogram is a cue-conflict stimulus: Texture cues specify an unbroken, unslanted surface, in conflict with any variation in depth specified by binocular disparity. We introduce a new visual texture—the starry night texture (SNT)—that is incapable of conveying slant, depth edges, or texture boundaries, in a single view. For SNT, changing density is equivalent to changing intensity, so an instance of the texture is characterized (up to the random locations of the texture elements) by what we call its densintensity. We consider deviations from the ideal that are needed to realize the texture in practice. In three experiments with computer-generated stimuli we examined human perception of SNT to show that (1) the deviations from the ideal that were needed to realize SNT do not affect the invariance of its appearance across changes in distance of several orders of magnitude; (2) as predicted, observers match SNT across changes in distance better than other textures; and (3) the use of SNT in a slant perception experiment did not reliably increase observers' reliance on stereoscopic slant cues, as compared with the sparse random-dot displays that have been commonly employed to study human perception of shape from binocular disparity and motion. © 2004 Optical Society of America

OCIS codes: 330.5000, 330.1400.

1. INTRODUCTION

Consider a Julesz random-dot stereogram[1] (RDS), printed on an unwrinkled sheet of paper. The stereogram is composed of a rectangular array of pixels, each of which is white or black. The property that made this stimulus interesting, according to Julesz, was its being "devoid of all depth and familiarity cues except disparity." However, a modern Bayesian would conclude that the RDS is a cue-conflict stimulus. Real depth edges are almost always accompanied by a discontinuity in visual texture, and surface slants are almost always accompanied by texture gradients in the image[2,3]; thus texture cues specify flatness in a Julesz RDS. If head movements are allowed, then motion parallax cues also specify flatness. Julesz's observation that binocular disparity *per se* is sufficient to evoke depth is actually more impressive when one realizes the nature of these conflicts. However, the stated goal of creating a stimulus devoid of nonstereo cues was not in fact realized in the Julesz RDS.

Visual texture (or surface pattern) reveals information about depth because visual textures tend to be homogeneous. For our purposes, a homogeneous texture will be a texture with measurable statistical properties that are translation invariant across the surface. Projecting surface(s) into an image plane causes depth-dependent departures from homogeneity in the image, from which depth information can be recovered, under the assumption that the original texture was homogeneous.[4] In addition, if the texture contains texture elements of an identifiable shape, the elements may be cut or deformed in the image, as a function of changes in depth.

Thus a texture must not support the measurement of departures from homogeneity, nor contain visibly shaped texture elements, if it is to completely hide slant and edges in depth. These considerations dictate that the texture must comprise independently (randomly) placed points. And indeed, small, sparsely and randomly placed dots are often used to create stereograms when it is important to minimize single-image depth cues. Yet the sparse RDS is still not a perfect answer: Both slant and distance increase the density of the dots within the image, so even small sparse random dots convey some information about depth. In addition, the images of small luminous dots vary with distance, giving rise to a "proximity luminance" cue.[5] To completely hide changes in depth, one would also have to create a random-dot texture in which neither dot density nor dot intensity provides information about depth.

The starry night texture (SNT) is a random-dot texture that, ideally, renders depth edges and changes in surface slant invisible, under the laws of projection that normally govern the formation of images. SNT is by its nature unusual, and it is not clear that the human visual system would know, on first being introduced to it, that texture cues in the image are indeed useless; one interpretation of experiment 3 is that observers treat images of SNT as though they contain depth information even though they do not. Regardless of how the human visual system interprets images of SNT, part of our intent is simply to point out that there is, in fact, a texture that can hide changes in depth remarkably well. For this reason, we describe the texture not in terms of the images it creates

but rather in terms of its physical instantiation as a collection of infinitesimal luminous disks sitting on an otherwise black surface in space.

In this paper we first introduce the ideal SNT and explain in greater detail why it hides depth information. We then describe approaches for handling two problems in implementing SNT: Images of ideal SNT have uniformly infinite luminance, so we restrict the number of dim stars in the texture; and physical display devices are limited in their ability to portray small, high-intensity light sources, so we discard instances of the texture that give rise to very bright stars in the image. Finally, we describe three experiments with SNT that test (a) whether our deviations from the ideal compromise the ability of SNT to hide changes in depth (to a large extent they do not); (b) whether, on account of its invariance in appearance across changes in distance, SNT is more easily matched across changes in distance than are other textures (it is); and (c) whether there is any evidence that the visual system gives greater credence to stereo depth cues when texture cues have been eliminated through the use of SNT, as compared with other textures (we find little evidence for this).

2. THEORY AND IMPLEMENTATION

We begin with definitions.

- An infinitesimal flat star (henceforth, a star) is a flat luminous disk, so small that its shape cannot be seen, that emits light from one of its sides. Its luminous intensity L in the direction ϕ relative to the surface normal of the star is $L = I\cos(\phi)/\pi$ candelas (cd), where I is luminous flux in lumens (lm).
- A starry night texture in the world consists of randomly and uniformly distributed stars, tangent to a black surface, with flux values $\{I\}$ that are distributed as $1/I^2$.
- An instance of SNT is fully characterized by two parameters: a lower bound, $\alpha > 0$ on star flux, and densintensity D. D is the number of stars per square meter expected to have luminous flux greater than 1 lm.
- A starry night image depicts SNT-textured surfaces. The image can be represented as a new set of stars on an image plane, where each star in the world is replaced by a star in the image, such that the direction and apparent brightness of each star is unchanged at the observer. In fact, the image of SNT is itself an SNT (Subsection 2.D), so in many cases, images of SNT can be created simply by assigning fluxes (according to the $1/I^2$ distribution) to stars in the image rather than by computing the image of each star from the three-dimensional scene.

For comparison, we consider three other textures besides SNT: lattice, random, and random flux. All textures comprise stars against a black matte background. In the lattice texture, stars are of constant flux and spatially arranged in a regular lattice. In the random texture, stars also exhibit constant flux, but their locations are uniformly distributed by area. In the random-flux texture, the distribution of flux among the stars is uniform (within some finite flux interval), and the locations of stars are uniformly distributed by area. In Fig. 1 these textures are shown. Note, however, that we have

not compensated for nonlinearities of the paper reflectance, so the random-flux texture and SNT are illustrative rather than accurate depictions.

A. Ideal Starry Night Texture
The ideal SNT is defined as follows:

1. Stars are independently and uniformly distributed by area on a black and matte surface. This prevents pattern matching between patches of the surface and, in conjunction with items (2) and (3) below, guarantees uniform distribution of dots within the image.
2. The distribution of the flux values of the stars is described by the probability density function: $D(I) = 1/(hI^2)$, where $I \geqslant \alpha$ and $h = \int_\alpha^\infty D(I)dI$ is a normalizing constant. This describes stochastically the distribution of flux among the stars in the SNT. It does not specify the actual number that will occur per unit area. In the case of the ideal SNT, α is chosen arbitrarily close to 0, and the SNT is characterized by just one number, its densintensity. A consequence (see below) is that the expected number of stars, per unit area in the image, exceeding some intensity value, is invariant to changes in the distance of the imaged surface.
3. SNT is composed of stars that are infinitesimal in size, so that dot size, shape, and partial occlusion of texture elements cannot provide information about depth structure within the scene. Because the stars are sufficiently small, the point-spread function of the eye (or other imaging device) makes a bright distant star indistinguishable from a dim near one.

Because stars are flat-disk Lambertian emitters, the amount of light reaching the image from a given star is proportional to the cosine of the local slant of the surface.

The ability of SNT to conceal information about the distances and slants of surfaces originates from the fact that changes in distance and slant do not cause any change in the statistics of the image. With other textures, statistical changes that result from changes in distance and slant are a consequence of homogeneity[6] in the texture, which yields a usable depth cue.[7] Of course, knowledge of a particular instance of the SNT does permit an estimate of slant and distance. However, knowledge of the statistical properties of SNT is insufficient for the inference of anything about the distance, slant, or edge locations of a depicted surface textured with the SNT. This is illustrated in Fig. 2.

Proof that images of SNT are devoid of depth cues. To provide an insight as to why images of SNT are devoid of depth cues, we show that this is true for a small surface patch at the center of the field of view. The full proof can be found in Zabulis and Backus.[8] We will show that the statistical properties of an image region w that depicts a SNT-textured surface S are invariant to the distance and slant of this surface. It suffices to show (1) that w is textured by SNT and (2) that the SNT of w has densintensity D, same as S.

To establish item (1), we must show that the x and y image coordinates of the stars in w are distributed uniformly and independently and that their intensities are distributed as $1/I^2$. First, note that the spatial layout of the stars in w is the same as in S after two scaling opera-
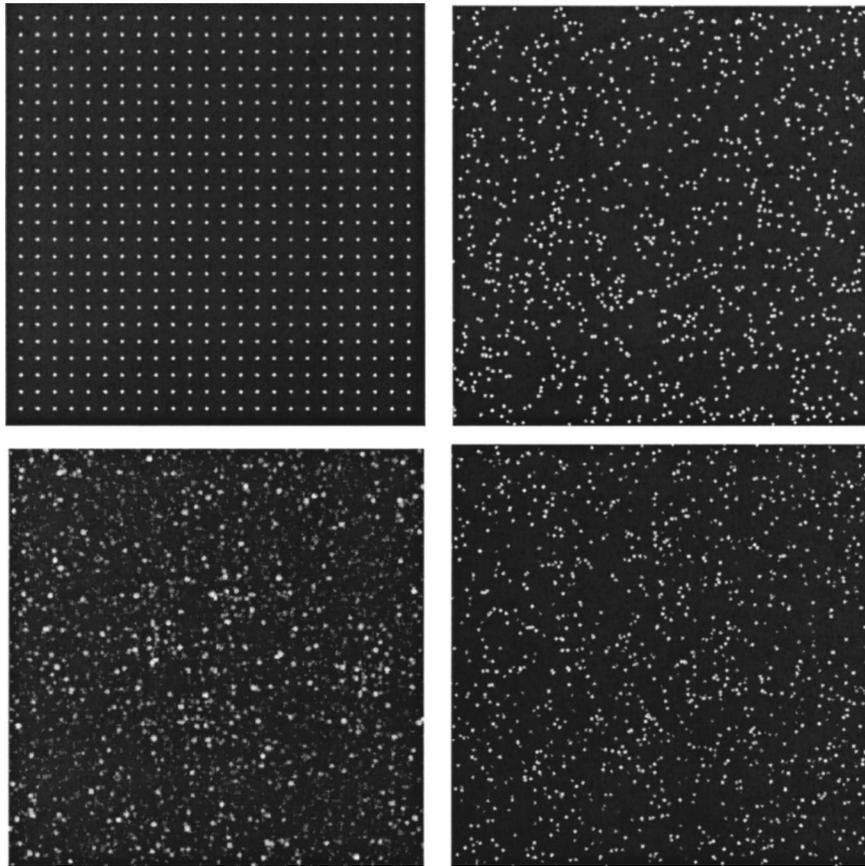
Fig. 1.　Texture samples (top left, clockwise):　lattice, random, random flux, and SNT (rendered with the algorithm of Subsection 2.D).

tions:　uniform scaling of linear dimensions by a factor of $1/d$, where d is the distance to S, and scaling by $\cos(s)$ in the tilt direction of S, where s is the slant of S.　Thus the stars in w have spatial positions that are related to those of the corresponding stars in S by an affine transformation.　Since the x and y coordinates of star spatial positions in S are distributed independently, so too are x and y coordinates of the stars in w.[9]　From the definitions of SNT and the imaging system (that replaces stars in the world by identical-appearing stars in the image), it follows that, as the diameter of w tends to zero, the flux of each star in w becomes equal to the flux of the corresponding star in S, times the factor $k = \cos(s)/d^2$.　Thus the flux of every star in w is equal to the flux of the corresponding star in S multiplied by a fixed scalar, k.　It follows that the fluxes of the stars in w are distributed as $1/I^2$, so w is SNT textured.

　　To show item (2), consider that the expected number of stars in S, per square meter, exhibiting a luminance 1 lm or greater, is D.　The area of w is k times the area of S. The flux of each star in w is k times the flux of each corresponding star in S.　Thus the expected number of stars in w with a flux of 1 lm or greater is $(\int_1^\infty kDI^{-2}\mathrm{d}I)/k = D$ (k in the integrand compensates for change of star flux from world to image; k in the denominator compensates for change in area).　Thus w has densintensity D, and image statistics are invariant to changes in distance and slant.

B.　Theoretically Realizable Starry Night Texture
The ideal SNT is impossible to realize for two reasons, one theoretical and one practical.　The theoretical reason is that distribution as $1/I^2$ for flux in SNT implies an infinite number of stars as α, the lower bound on I, tends to 0.　As with Olbers's paradox in astronomy, the combined luminance from these stars would cause an infinitely luminous image.　Thus the light coming from stars at the lower end of the flux distribution must be limited; we chose to effect this limit by choosing α positive.　Densintensity now determines the expected (finite) number of stars per unit area, in addition to being equal to the expected number of stars, per square meter, of intensity greater than 1 lm.[10]　Interestingly, although any instance of SNT has finite luminance, expected mean luminance is still infinite owing to rare, arbitrarily bright stars: $\int_\alpha^\infty (D/I^2)I\mathrm{d}I = \infty$.　Thus mean luminance is not a useful way to characterize SNT.

C.　Practically Realizable Starry Night Texture
Physical displays are limited in their ability to portray very bright stars.　Thus the distribution of star fluxes must be tailored so that no star is too intense.　This can be achieved by imposing an upper limit β on the probability density function of star flux so that it becomes $D(I) = \tau/I^2$ within the interval $\alpha < I < \beta$ and 0 elsewhere; τ is equal to $\alpha\beta/(\beta - \alpha)$.　As a practical matter, the densintensity D must be chosen in light of β.　One would

like for the actual star fluxes to have plausibly come from a $1/I^2$ distribution; but if the densintensity is too large, then the presence of stars with flux near (but below) β, together with the absence of stars with flux greater than β, would allow the observer to infer that the imaged texture is not SNT.

Depth cues in computer images of SNT. Besides texture statistics, there exist other cues to depth in computer-generated images of SNT. Ideally, stars would be depicted by dots small enough that, given their distance from the eye and its point-spread function, the area they occupy on the retina is indistinguishable from the area occupied by the image of an equally intense point light source. In our experiments with SNT we used CRTs, so we were unable to render dots that were simultaneously very small and very bright, and dot sizes were visible as a result.

It is a consequence of the fact that $\alpha > 0$ that, as the distance between the surface to the center of projection decreases, there will eventually be completely dark regions between visible stars in the image. When distances are greater, the dimmest stars are not individually discriminable, but they still contribute to image luminance. We refer to the luminance in the image between two discriminable stars as background luminance. Background luminance is potentially a distance cue; however,

experiment 1 showed that humans are unable to make use of this cue even for hundredfold changes in distance.

D. Stimulus Formation

Since the image of SNT is itself an SNT, one approach to creating images of SNT surfaces is to use SNT directly instead of computing the image as a projection of the surface. In that case, however, the lower cutoff α will describe a bound in the image, not in the world. Therefore we also created images by projecting texture in the world onto the image plane. A three-dimensional model of the surfaces was created (as a computer representation), thereby defining the distance and slant of each point on the surface. Stars, with fluxes drawn from the $1/I^2$ distribution with $I > \alpha$, were then added to the surfaces until the desired densintensity was reached. Any of these stars that were unoccluded to the observer were then added to the image. It can be shown[8] that the luminous flux I_{image} of a star in the image, to make it appear the same as the corresponding star in the world, is $I_{\text{image}} = I_{\text{world}} \cos(\phi)(r_{\text{image}}/r_{\text{world}})^2/\cos^3(\theta)$, where I_{world} is the flux of the star in the world, r_{world} is the radial distance from the observer to the star, r_{image} is the distance from the observer to the image plane, ϕ is the local slant of the surface relative to the line of sight, and θ is the eccentric-
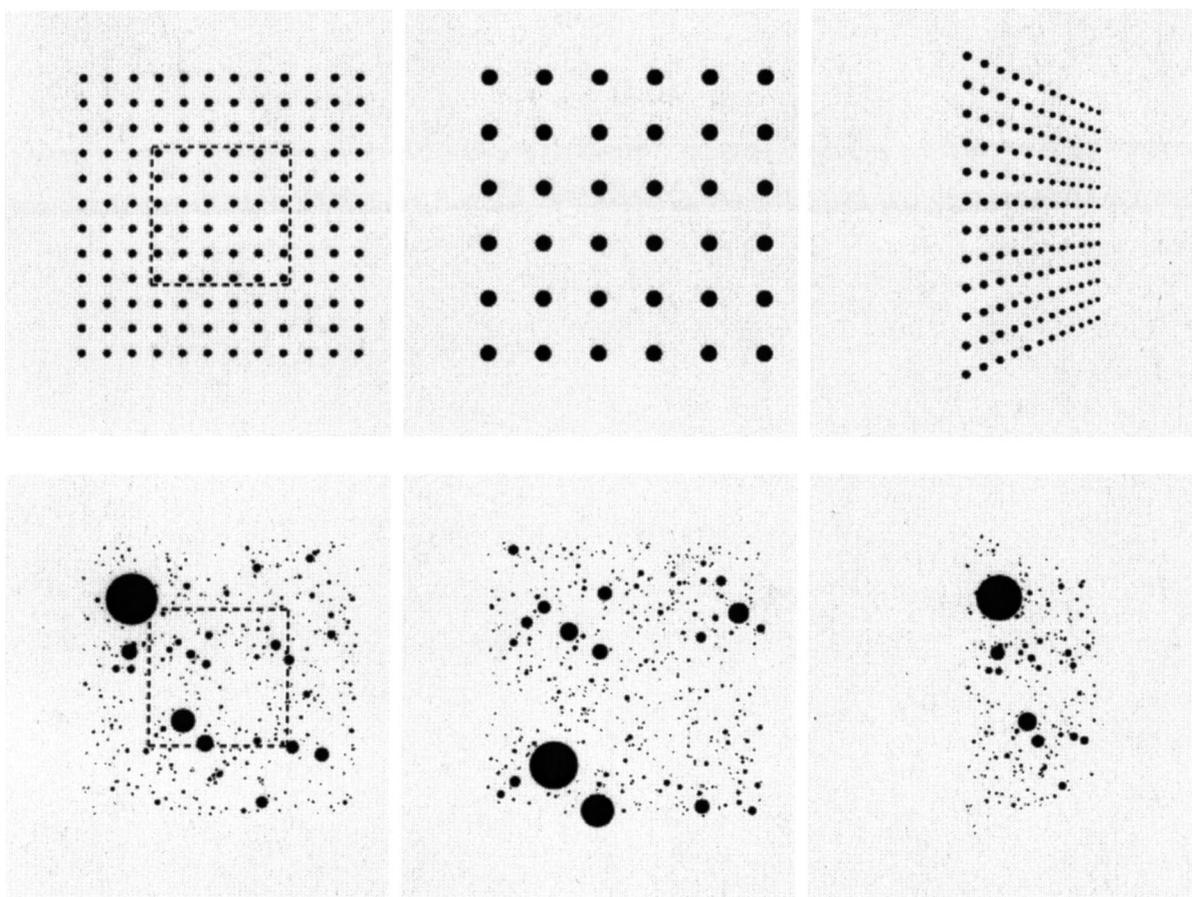

Fig. 2. Most textures give rise to depth cues, but SNT does not. Top row: A homogeneous and isotropic texture, on a planar surface patch, is imaged in three positions (left to right)—frontoparallel at distance L, frontoparallel at distance $L/2$, and slanted 60° at distance L. Circle areas indicate star intensity (not shape). The dashed square in the left panel shows the portion of the surface that is visible in the middle panel. Bottom row: same for SNT.

ity of the star (i.e., the size of the angle formed by rays from the observer to the closest part of the image and to the star).

To render the image, we added each star independently to the image as a 2×2-pixel antialiased dot. The resulting image was convolved with a raised-cosine kernel to scale overall luminance and give the dots a consistent round shape. A sample output is shown in Fig. 1.

3. EXPERIMENTS

Three experiments were conducted to characterize specific aspects of human perception for images of SNT. First, we were interested to know the perceptual consequences of truncating the flux distribution. Experiment 1 therefore compared distance-discrimination thresholds for simulated surfaces textured with SNT with thresholds for three other textures (lattice, random, and random flux). To anticipate the result, we found no visible differences between the images of near and far SNT surfaces, for a 20-fold change in distance; in contrast, a 10% or smaller change in distance sufficed for the other textures. Second, since texture is a cue to distance and slant, the visual system must also be capable of estimating the properties of textures in the world. In other words, the visual system should exhibit texture constancy: the ability to determine whether textures on surfaces at different distances are the same. If observers can make estimates of this sort for SNT, then they ought to see two SNTs at different distances as having equal densintensity if and only if the densities of their *images* are equal. In experiment 2 we asked observers to match textures at different distances, and we found large effects of simulated distance for the random and random-flux textures but not for SNT. Third, our interest in SNT derived in part from its ability to hide information about distance and slant, as this would allow stereo and motion depth cues to be studied in isolation. Is it the case that SNT differs from traditional sparse RDS, causing the visual system to disregard pictorial cues, relative to binocular disparity? In experiment 3 we asked observers to estimate the slant of surfaces that were defined by disparity. For SNT, texture cues were uninformative, whereas for other textures, texture cues indicated zero slant. We found that slant was underestimated for SNT and random textures to similar extents. This implied that accommodation by the lens of the eye, or other depth cues, was more important than dot density in explaining the underestimation of surface slant for our stimuli.

A. Apparatus

The stimuli were displayed on a haploscope consisting of two 19-in. (1 in. = 2.54 cm.) Clinton medical monochrome CRT displays, each seen in a mirror by one eye. These images were displayed at a resolution of 1280 $\times$ 1024 pixels and at a refresh rate of 75 Hz, at an optical distance of 42 cm from the eye. In experiments 1 and 2, stimuli were monocular so only the right CRT was active, and observers wore a patch over the left eye. The mirrors were 6×10 cm, rounded in shape to fit close to the eye, and oriented at 45° relative to the line of sight. The face of each CRT was always perpendicular to the line of sight from the eye to the center of the screen. The CRTs were calibrated in luminance and spatially, as in Ref. 11. The calibrated area was $48° \times 36°$ on each monitor. The room was completely dark except for scattered light from the displays. A computer generated the stimuli before the experiment and displayed them during the experiment, using Matlab and the Psychophysics Toolbox.[12] Observers' responses were made with a numeric keypad and collected by the computer.

B. Experiment 1

The purpose of this experiment was to determine the range of distances over which the appearance of implemented SNT is invariant to human observers and to compare this range with the corresponding ranges for the lattice, random, and random-flux textures.

1. Stimuli

Viewing was monocular. Stimuli depicted a frontoparallel background surface. Half of the stimuli depicted a foreground surface as well, which occluded the central region of the background surface. The foreground and background had the same texture.

Figure 3 (top) illustrates the scene. The background surface was a 44 cm $\times$ 44 cm frontoparallel, flat, square patch at a simulated distance of 1 m from the observer. In one of the two stimuli presented in a given trial, a foreground surface was also present. This surface was also frontoparallel and was closer than the background, thus occluding the central one fourth of the background square. For the lattice texture, the position of the texture on the foreground patch was centered so as to be a pure expansion relative to the center of the patch. The dot diameter was 0.03° (raised-cosine luminance profile).

The texture density ranged from just dense enough to do the task to a density at which almost all pixels in the image were occupied by some part of a luminous dot (see abscissa in Fig. 3's graph). The star flux for the lattice and random textures was 7.5×10^{-6} lm. For the random-flux texture and SNT, flux was tuned separately for each density, so that a maximal number of stars were visible: 3.7×10^{-9} to 7.5×10^{-6} lm for the random-flux texture, and $\alpha = 3.7 \times 10^{-9}$ to $\beta = 4.5 \times 10^{-5}$ lm for SNT.[13]

2. Task

On each trial, two stimuli were presented for 2 s each, separated by 1 s of black screen. One stimulus contained the foreground patch in addition to the background patch. The observer's task was to indicate which interval contained the foreground patch by using a key press. Visual feedback was given after each trial. A three-correct, one-wrong staircase procedure controlled the depth interval between the background and the foreground patches. Novel instances of textures were used on each presentation. The staircase terminated after 15 reversals; threshold was the average of the last 12 reversals. This yielded an estimate of a 79% correct detection threshold. Each such estimate was one independent measurement. Observers rested as needed.
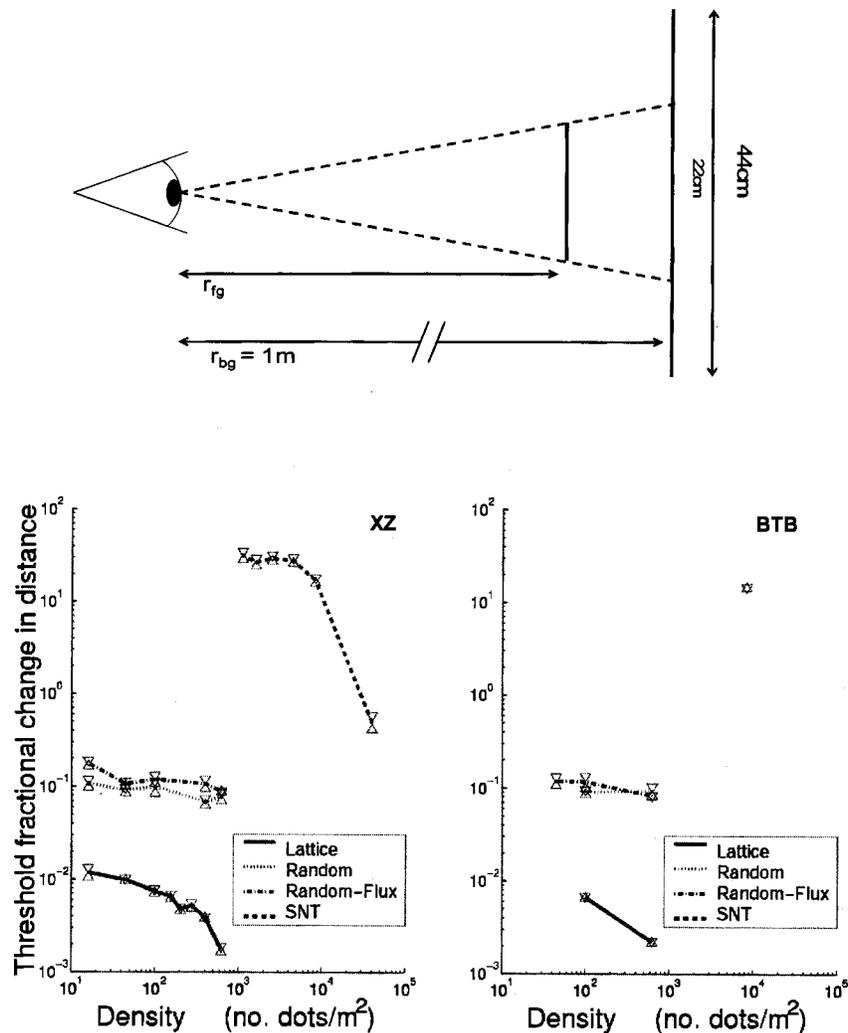
Fig. 3. Top: Schematic diagram of the scene simulated in experiment 1 (side or top view). r_{fg} was the staircase variable (distance to the foreground object), and r_{bg} was 1 m (distance to the background object). Bottom: Data from experiment 1, for two observers XZ (left) and BTB (right). The graph plots the threshold fractional change in distance ($r_{bg}/r_{fg} - 1$) required to identify which of two stimuli contained an occluding surface patch, for four textures, as a function of dot density. Data points are the average of three threshold measurements from three separate staircase procedures. The centers of the triangles mark the standard errors of the mean.

3. Results

Figure 3 (bottom) plots the change-in-distance threshold for each of the four textures for two observers (the authors). The abscissa is the texture density d_W, and the ordinate is the fractional change in distance between the background and the foreground, required for discrimination. The fractional change is $r_{bg}/r_{fg} - 1$, where r_{bg} and r_{fg} are the distances from the observer to the background and foreground, respectively.

For both observers, distance discrimination was easiest with the lattice texture; a change in distance of approximately 1% was sufficient to identify the foreground-containing stimulus. For the random and random-flux textures, a roughly 10% change in distance was required. For SNT, a roughly 20-fold change in distance was required. Increasing density caused a decrease in threshold; this effect was most pronounced for the lattice texture and SNT.

4. Discussion

The main observation from this experiment is that the monocular depth cue from texture, as measured by the

change in distance required for one to see a change in the image, was 2 orders of magnitude weaker for SNT than for the random and random-flux textures, and 3 orders of magnitude weaker than for the lattice texture. Thus depth is much more difficult to see in scenes that are textured with SNT than other textures, and the cutoffs for star flux, which were needed to realize SNT, did not cause failure of the theoretical ability of SNT to hide changes in distance.

For the lattice, random, and random-flux textures, observers felt discrimination to be based on the fact that stars in the foreground surface appeared sparser and brighter than stars in the background surface. As texture density increased, an additional cue became available in the lattice texture: At the corner of the foreground surface, the spatial arrangement of the three background dots and one foreground dot deviated from square when the foreground surface was present. For SNT, distance discrimination seemed to be based on the apparent darkness of the interstar regions: When the foreground surface was close enough to the observer, it be-

gan to run out of the dim dots that contributed to background luminance. At the highest density (for observer XZ), the background was brighter, and it seemed that this made it easier to see decrements in brightness for the near surface.

It is interesting to note that the random-flux texture did very little better than the random texture at hiding changes in distance. Randomizing intensity was not sufficient to hide changes in distance; the particular $1/I^2$ distribution of star fluxes in SNT is therefore a critical factor in its ability to hide changes in distance.

C. Experiment 2

For most textures, texture in the world can be inferred from the image, together with estimates of distance and slant. If the perceived distance is made to increase while the image stays the same, one would expect the apparent size of the texture (i.e., its element size and spacing) to increase.[14] It is therefore of interest to ask whether the apparent surface pattern for SNT also changes as a function of distance—in principle, it might not because the visual system might correctly infer that the statistics of the pattern on the surface are the same as in the image, no matter what the distance to the surface. In experiment 2 we measured this indirectly by using a nulling paradigm, in which observers adjusted the density (or, for SNT, the densintensity) of a texture at a variety of distances, until it matched the texture at a standard distance. Observers were instructed to match texture properties in the simulated world, not the image.

For SNT, this matching task can be done accurately by matching density in the images of the two surfaces, even when distance is not correctly estimated. Thus we predicted that SNTs would be adjusted until the densintensity of the test surface (in the world) matched that of the reference surface (in the world), independent of the simulated distance of the test. For other textures, perfect distance constancy would make the same prediction: Texture would be adjusted in the world until it was correctly matched between the test and the reference. However, perceived distance typically varies by less than the variation in simulated distance. This means that we predict that observers will be systematically inaccurate when matching textures other than SNT across changes in simulated distance (relative to the nominally correct responses predicted by simulated distance). If simulated distance is halved, for example, perceived distance might be reduced by only 25%; thus the observer would adjust density in the test surface until its image had a density that was 0.75^2 times that of the reference surface's image rather than the nominally correct amount of 0.5^2. At that point, its density in the simulated world would have been adjusted to $4(0.75)^2 = 2.25$ times that of the reference surface. We therefore predicted that observers' density settings (in the world) for random and random-flux stimuli would vary as a function of distance. That is, observers would compensate for distance but by an amount smaller than predicted from our distance manipulation.

1. Stimuli

Viewing was monocular. Stimuli depicted two square surface patches of equal size, the test surface and the ref-

erence surface. Both were facing the observer and resting on a checkerboard floor. The reference surface occurred always in the same position, whereas the test surface varied in distance from the observer. In all stimuli, both surfaces were textured with the same type of texture. The scene geometry and a sample stimulus are shown in Fig. 4.

The reference surface was positioned at a distance of 20 m from the observer, on the right side of straight ahead. The test surface was on the left, and its distance took on one of 21 values logarithmically spaced between 10 and 40 m. The size of the surfaces, in the simulated world, was 2.2 m × 2.2 m, so their angular subtense varied from 3.2° to 12.6°. The observer was 1.25 m above the checkerboard floor, which extended nearly to a horizon at eye level. The checks on the floor were gray (luminance 42 lm/m² in the image) and black, and the sky was black.

Each star on the surface was portrayed by an antialiased dot that occupied 2 × 2 pixels (0.19°) for all simulated distances. Densities and fluxes (in the simulated world) were, respectively, random texture: 240 dots/m² and 2.6×10^{-3} lm; random-flux texture: 240 dots/m² and 5.2×10^{-6} to 3.2×10^{-3} lm; and SNT: 3845 dots/m² with $\alpha = 5.2 \times 10^{-6}$ and $\beta = 31.2 \times 10^{-3}$ lm. Dot density in the test patch was adjusted by the observer and was the dependent variable.
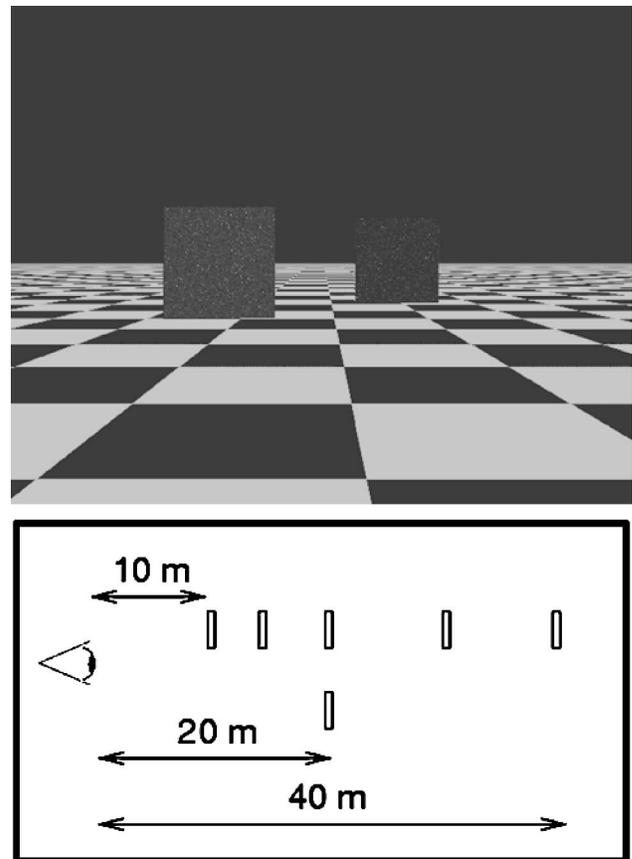

Fig. 4. Top: Schematic diagram of scenes used in experiment 2 (top view). The eye is shown at the left. The reference surface was at a fixed distance of 20 m from the observer, and the test surface varied in distance from 10 to 40 m; some of those positions are shown in the figure. Bottom: Sample stimulus as seen by the observer during the experiment. The test surface is at a simulated distance of 15 m.
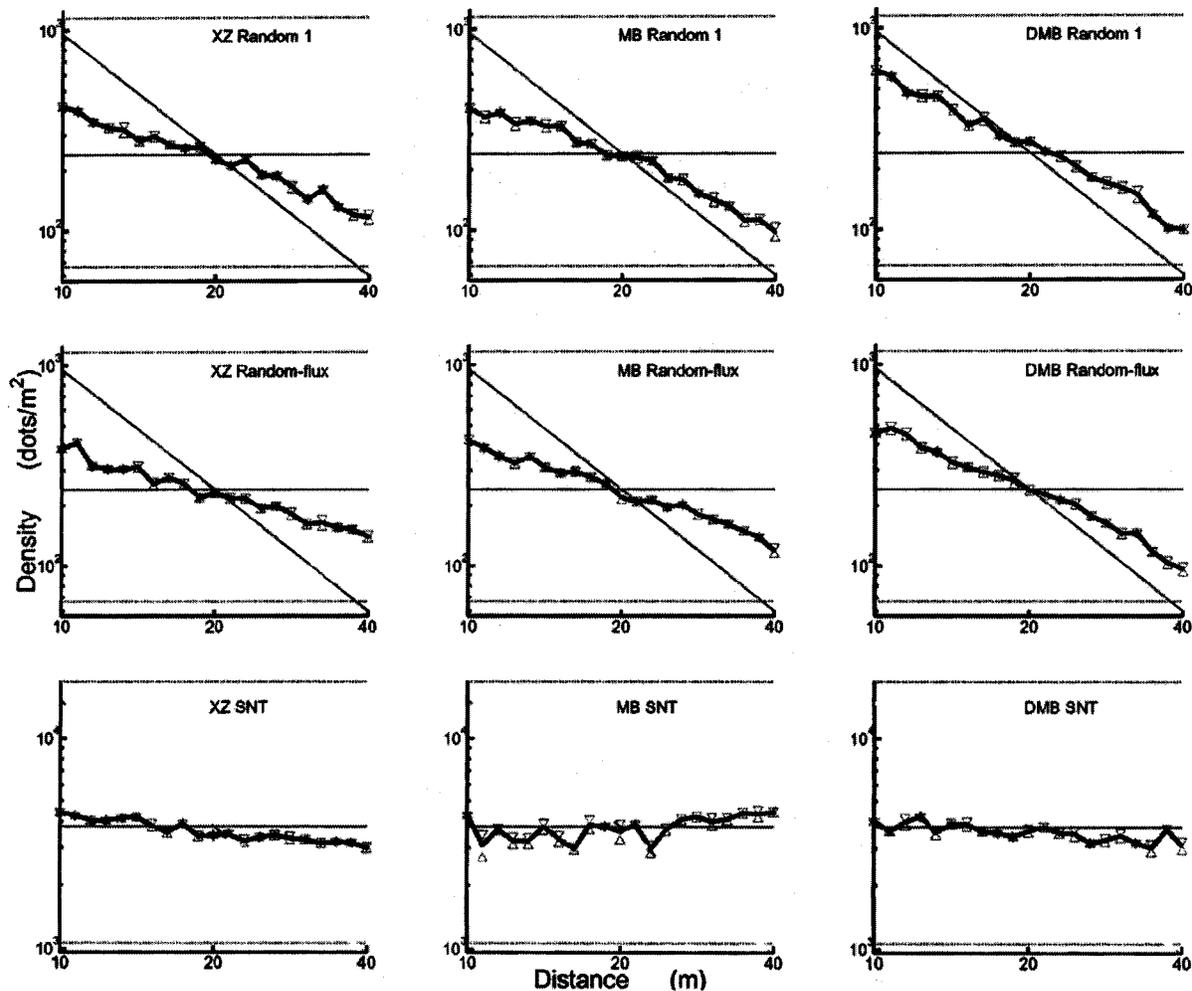
Fig. 5.   Data from conditions random 1, random flux, and SNT of experiment 2.   Columns are data from the three observers (left to right: XZ, MB, and DMB).   Rows show data for different conditions (top to bottom: random 1, random flux, and SNT).   The abscissa is the distance of the test surface from the observer.   The ordinate is the observer's setting of density for the test surface.   Data points are from three independent trials; triangles mark standard errors.   Dotted horizontal lines show the range of allowed responses.   Solid horizontal lines plot the density (in the world) of the reference surface, which is therefore the setting predicted by complete perceptual constancy.   Diagonal lines plot the settings predicted if observers simply matched dot density in the image of the test surface to dot density in the image of the reference surface.

In these stimuli, the cues available to the observer to indicate the ratio of distances to the test and reference surfaces were texture gradient and linear perspective in the checkerboard floor, relative sizes of the surfaces in the image (under the assumption they were the same size in the world), and declinations in the lines of sight to the bottoms of the surfaces.   When the random and random-flux textures were used, star flux in the image was an additional cue.

2. Task

Observers adjusted the density of the texture on the test surface by using key presses until the test surface appeared to have the same texture as the reference surface. Observers were explicitly instructed to equalize texture in the world, not in the image.   They took as long as needed to make a careful match and rested as needed.

The experiment had four conditions, blocked by session.   In each condition, one type of texture was used. The conditions were random 1, random flux, SNT, and random 2.   In the first three conditions, the observer adjusted only density—the fluxes of the stars in the image were clamped at their correct values, given the simulated distance.   In condition random 2, observers adjusted both density and flux.

Approximately 400 levels of each property were available; step size was less than the standard deviation in the observers' settings, and the ranges extended well below and above the observers' final responses.

3. Results

Three observers participated in the experiment: XZ, MB, and DMB; the first was one of the authors, and the other two were naïve to the experimental hypotheses.   Figure 5 shows the results from the first three conditions.   Of the nine graphs, the three columns represent different observers, and the three rows represent the different conditions.   In each graph the abscissa is the distance from the observer to the test surface.   The ordinate is the density of texture on the test surface (d_W), set by the observer. Central horizontal lines plot the density of the texture on the reference surface, indicating the response predicted

by complete perceptual constancy. Diagonal lines indicate the predicted settings for complete failure of constancy by matching density in the images. Dotted lines show the limits of the range of densities that was available to the observers.

The data for condition random 2 are shown in Fig. 6. Each column represents the responses of one observer. Graph arrangement is as in Fig. 5. In the bottom row of graphs, the abscissa is distance, and the ordinate is star flux. Diagonal, horizontal, and dotted lines have the same meanings as in the top row; they refer to density in the top row and to flux in the bottom row. In all graphs, each data point shows the average of three independent setting pairs (density, flux), and centers of the triangles indicate standard errors.

4. Discussion

The extent of compensation for simulated distance in experiment 2 can be summarized by the slopes of the best-fitting density-versus-distance lines in the log–log plots in Figs. 5 and 6. A slope of zero indicates complete compensation, so the density (in the world) on the test surface was set equal to the density (in the world) of the reference surface. A slope of -2 indicates no compensation because this describes the case where the images have equal density.

Figure 7 plots these slopes for each of the three observers in each of the four conditions in a bar graph. The data are remarkably consistent across observers. They show that density in the test surface was adjusted nearly correctly for SNT and that simulated distance was undercompensated for in conditions random flux, random 1, and random 2.

The density settings for the random and random-flux textures (Fig. 6) show that observers took the simulated distance into account for purposes of equating texture because images of far test surfaces were set to be denser

than images of near test surfaces, in agreement with the geometry of the simulated scene. However, the magnitude of this effect was smaller than predicted by complete constancy. This shortfall is not surprising because the perceptual consequences of manipulating simulated distance in computer displays are usually smaller than the manipulation predicts.[15] Thus one can interpret the data for the random and random-flux textures as showing that observers are able to accurately estimate the densities of textures in the world by using an internal estimate of distance, but that in our experiments this internal estimate varied less than the simulated distance.

Figure 7 also shows that observers' responses were different for SNT compared with their responses for the other textures. The data show that settings for SNT, measured on the test surface in the world, varied little with changes in simulated distance. This is not surprising, assuming accurate distance constancy. Settings for SNT, but not the other textures, would remain unaffected by a discrepancy between simulated and perceived distances. Of course, simply matching the image of the test surface to the image of the reference surface would give a similar result. Did observers ignore distance when matching SNT? Or did they take distance into account and correctly infer that the image of SNT ought not to change across changes of distance?

Observers' density settings for SNT were not completely invariant to changes in simulated distance. Observer MB adjusted densintensity up with increasing distance, whereas observers XZ and DMB adjusted densintensity down with increasing distance. This is what observers would do if they were estimating what the surface would look like at different distances and doing so slightly incorrectly.

One might be concerned by the fact that intensity was a distance cue in condition random 1 but not SNT: Perhaps dot intensity was the basis for density settings in
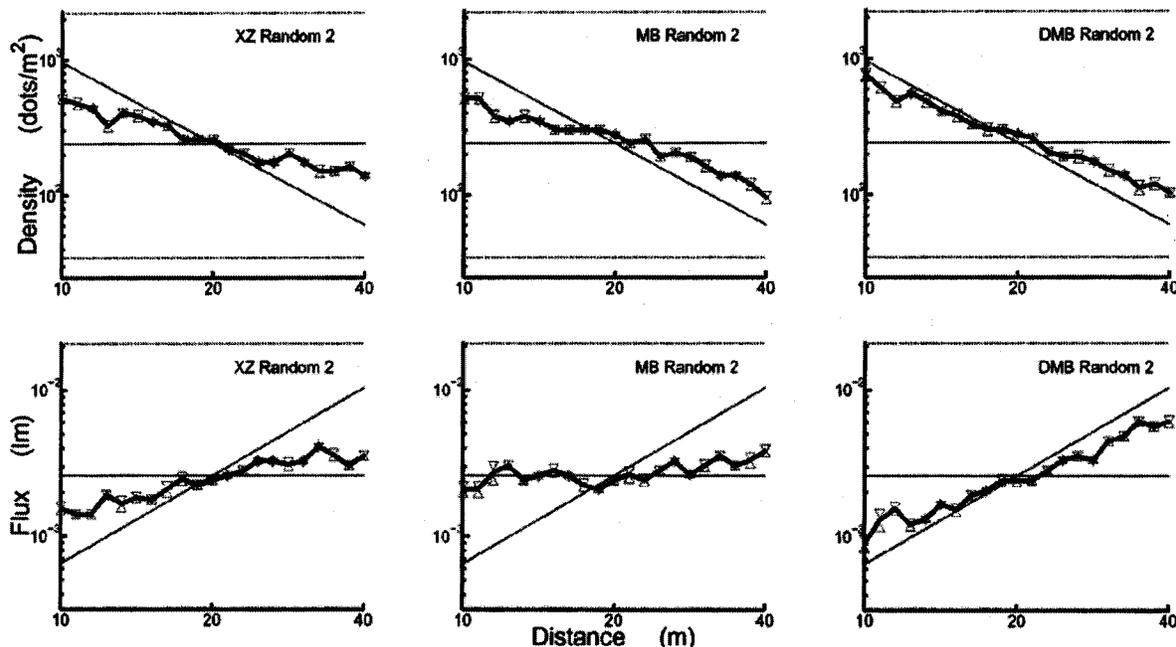

Fig. 6. Data collected for condition random 2 of experiment 2. Columns are data from the three observers. The top row shows density settings, and the bottom row shows flux settings. Graphs are as in Fig. 5.
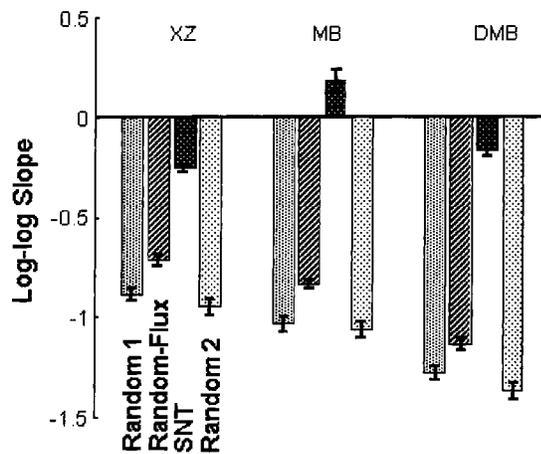
Fig. 7. Summary of data collected in experiment 2. Each bar plots the slope of the regression line in the log–log plot of density setting versus simulated distance, from the data in Figs. 5 and 6. Each group of four bars are data from one observer, showing the slopes for conditions random 1, random flux, SNT, and random 2. A slope of zero indicates that the texture density of the test surface was adjusted to be equal to the texture density of the reference surface, across simulated distances. A slope of -2 indicates that the texture density in the image of the test surface was adjusted to be the same as the texture density in the image of the reference surface, across simulated distance. Error bars depict one standard error on either side of the mean for the estimate of log–log slope.

condition random 1. However, in condition random 2, intensity had to be adjusted according to the distance cue provided by perspective. Observers' density settings were similar in conditions random 1 and random 2, so it seems likely that perceived distance was similar for SNT and the other conditions. We therefore conclude that observers scaled SNT for distance and that their visual systems correctly inferred that the image of SNT does not change when the distance to the surface changes.

D.   Experiment 3
In displays with real objects, perceived depth is consistent with binocular disparity cues.[16] In stereoscopic computer displays, however, perceived depth varies less than simulated depth.[15] This is presumably because nonstereo factors conflict with the stereoscopic cues in these displays and indicate the absence of depth modulation across the display. The nonstereo factors may include a bias (prior assumption) of equidistance to the elements in a scene,[17] but if this were a major cause of underappreciated depth in computer displays, one would expect it to affect displays with real objects as well. Computer displays do contain depth cues, including the absence of differential blur across the display,[18] the related cue of lack of change in accommodative demand across the display, failure of texture gradients and perspective to simulate distance and slant changes in the display, failure of the texture elements to change intensity and size with simulated distance, and failure of the texture elements to change aspect ratio with simulated slant.

Experiment 2 showed that, across changes of perceived distance, the visual system correctly infers that the image of SNT ought not to change. By the same token, the visual system might be able to infer that SNT contains no

useful information about distance; it might therefore give less weight to texture cues in computer displays when surfaces are textured with SNT, as compared with other textures. In experiment 3, we aimed to determine whether stereoscopically rendered slant is perceived as being greater in magnitude when displays were images of SNT than when they were images of other textures.

1.   Stimuli
Stimuli were stereo image pairs of slanted, elliptically shaped surface patches with lattice, random, random-flux, and SNT textures. Dots were as in experiment 2. The surface patches had approximately circular images and subtended 18° of visual angle. Binocular disparity was the only cue that was manipulated explicitly. Vergence and vertical disparity were consistent with a viewing distance of 1 m, and accommodative demand was determined by the distance to the screen (42 cm). Images were generated as follows: The surface was represented as frontoparallel (unslanted) in space. The locations of stars in the images were then determined by projection separately for the left and right eyes. Then the horizontal coordinates of the dots in the left and right images were scaled (relative to the center of the image) by factors $HSR^{1/2}$ and $HSR^{-1/2}$, respectively, where HSR was the desired horizontal size ratio between the left and the right eyes' images, as determined by solving for HSR in the slant equation $S = \tan^{-1}[(1/\mu)\ln(HSR)]$, where μ was the vergence angle of the eyes at the 1-m simulated viewing distance.[11] Thus the binocular disparity gradient was that of a flat surface with a slant of S deg at 1 m. The dot density gradient was not affected by this scaling operation, so it continued to specify a slant of 0°.

The slant specified by the stereo signal varied from $-60°$ to 60° in steps of 5°. A small fixation mark consisting of four vertically aligned dots was provided at the center of the disk, and observers were instructed to fixate this mark during the experiment. Stimuli were prepared taking into account interocular distance. All 25×4 stimuli were presented once, in a random order, during a single session. Dot density and flux were the same for each texture.

2.   Task
On each trial the observer viewed the slanted surface in alternation with a response stimulus. The observer controlled the alternation between the test stimulus and the response stimulus by using key presses. The response stimulus contained a pair of line segments,[19] and the angle between them was adjustable by use of key presses. Observers indicated the perceived slant by adjusting this angle. When the observer was content, he made a separate key press, and the computer recorded the slant setting and started the next trial.

3.   Results
Figure 8 (top) plots data collected from one of the observers. The four data series plot data for the four texture types. The effect of texture on the perceived slant was summarized by fitting a line to each data series; the slope of this line is the gain of the slant response relative to the
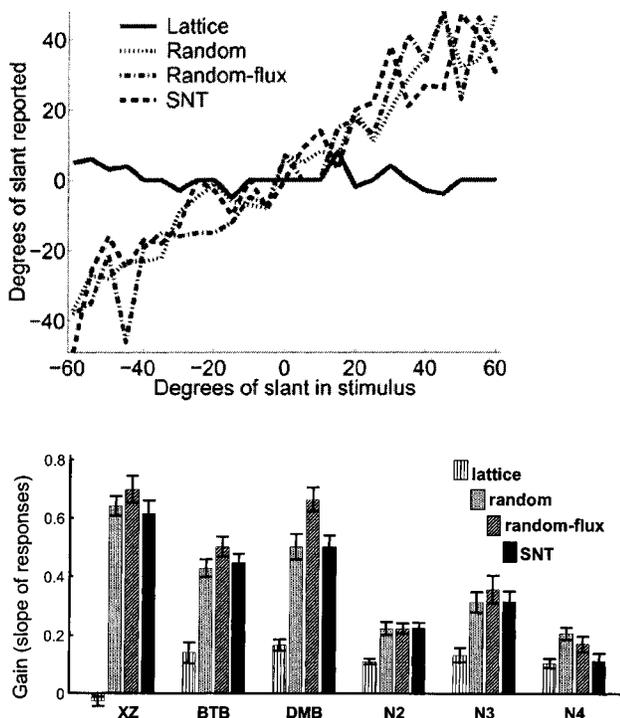
Fig. 8.   Top: Slant responses for one observer (XZ) in experiment 3.   The abscissa plots the slant in the stimulus that was specified by disparity cues.   The ordinate plots the observer's slant response.   Different curve styles correspond to different textures, as shown in the legend.   Bottom: Gain of the slant response, relative to slant specified by disparity.   Data from six observers, for four textures (lattice, random, random flux, and SNT).   Observers DMB and N2–N4 were naïve to the hypotheses of the experiment.   Each bar is the coefficient of regression for one data series in a graph of the form shown in the top panel. Error bars are standard errors.

slant specified by stereo cues.[19]   This stereo gain is plotted for all six observers in the bar graph of Fig. 8 (bottom).

4. Discussion

Observers' slant responses were smaller than the slant specified by disparity for all four textures.   Since we did not collect responses for real stimuli under full-cue conditions, we cannot say to what extent this reflected the mapping between stimulus and percept and to what extent the mapping between percept and response.[20]   However, it is clear that the lattice texture evoked very little perceived slant and that the other textures evoked perceived slants that depended in a similar way on the disparity gradient.

   Why was perceived slant so small for the lattice texture? From a Bayesian point of view, a square lattice in the image might be caused either by an unslanted square lattice in the scene or by a specific inhomogeneous texture, slanted by just the amount needed to create the lattice image.   The visual system evidently believed that the former scene was very much more likely to have caused such an image, as revealed by the (unslanted) percept it constructed—despite the fact that this percept gave very little weight to disparity.

   The random texture gives rise to a weak texture cue compared with the lattice texture, but the texture cue is

not completely absent, as it is for SNT.   We predicted that the visual system would interpret the texture gradient of zero in images of the random texture as indicating a slant of zero, causing perceived slant to be greater for SNT than for the random texture.   This prediction was incorrect.   It is possible that the visual system treats SNT as though it contains texture cues to slant or that other cues (circular outline, accommodation, dot size–shape) were more dominant than density as cues to flatness in the random texture[21–23] and that these other cues acted similarly for the random texture and for SNT.   In the latter case, it might be possible to reveal the predicted effect of greater perceived slant for SNT by increasing the distance to the CRT because the texture cue in the random texture would be given greater weight relative to stereo as distance increased[24] and because accommodation, dot size, and dot shape would become harder for the visual system to measure.

4. SUMMARY

The SNT is a novel texture that, in its ideal form, does not convey information about surface distance or slant when imaged.   To realize SNT, we truncated the distribution of star fluxes at both ends and used finite-size display elements to represent the stars.   Experiment 1 showed that the ability of SNT to hide changes in distance largely survived the truncations.   Experiment 2 showed that observers matched textures in the world across changes in perceived distance and that they appreciated that images of SNT are invariant to changes in distance.   Experiment 3 failed to show that binocular disparity is given greater weight by the visual system for SNT than for other textures, in a slant estimation task.   Whether SNT will be useful for its property of providing binocular disparity (or motion parallax) depth cues in isolation remains to be seen.   Perhaps with training, during which observers interact with real or simulated objects that are painted with SNT, the visual system might learn to appreciate that image statistics for SNT are invariant to changes in distance and slant.

ACKNOWLEDGMENTS

REFERENCES AND NOTES

1.   B. Julesz, "Binocular depth perception without familiarity cues," Science **145**, 356–362 (1964).
2.   J. J. Gibson, *The Perception of the Visual World* (Houghton Mifflin, Boston, Mass., 1950).

3. J. J. Gibson, *The Ecological Approach to Visual Perception* (Houghton Mifflin, Boston, Mass., 1979).

4. J. Malik and R. Rosenholtz, "Computing local surface orientation and shape from texture for curved surface," Int. J. Comput. Vision **23**, 149–168 (1997).

5. B. A. Dosher, G. Sperling, and S. Wurst, "Tradeoffs between stereopsis and proximity luminance covariance as determinants of perceived 3D structure," Vision Res. **26**, 973–990 (1986).

6. Homogeneity is sometimes distinguished from isotropy, with homogeneity referring to the regularity in spacing of textural elements and isotropy referring to regularity in the orientation of textural elements. SNT has stationary statistics across the surface, but this is the only sense in which it could be considered homogeneous or isotropic.

7. R. Rosenholtz and J. Malik, "Surface orientation from texture: isotropy or homogeneity (or both)?" Vision Res. **37**, 2283–2293 (1997).

8. X. Zabulis and B. T. Backus, "The starry night texture," Tech. Rep. IRCS-04-01 (Institute for Research in Cognitive Science, University of Pennsylvania, Philadelphia, Pa., 2004); ftp://ftp.cis.upenn.edu/pub/ircs/tr/04-01/.

9. This fact is easy to see for changes in distance that cause a uniform scaling of the image. But changes in slant cause scaling in a single direction, and the reader may suspect this would leave visible evidence in the form of an anisotropy. This is not the case. Points that are distributed randomly from a two-dimensional uniform distribution will remain so after a one-dimensional scaling, for if x and y are independent of each other across the set of points (x, y), then x and my will also be independent of each other across the set of points (x, my).

10. In principle, α might be greater than 1 lm. We therefore define densintensity as specifying the distribution from which the stars are drawn rather than the actual number of stars of intensity greater than 1 lm expected per square meter.

11. B. T. Backus, M. S. Banks, R. van Ee, and J. A. Crowell, "Horizontal and vertical disparity, eye position, and stereoscopic slant perception," Vision Res. **39**, 1143–1170 (1999).

12. D. H. Brainard, "The Psychophysics Toolbox," Spatial Vision **10**, 433–436 (1997).

13. The upper limit of the flux interval is an order of magnitude greater for SNT than for the other texture(s). However, few (if any) stars had this level of flux, owing to the sparseness of high-flux stars in SNT. In rare cases in which the resulting luminous intensity magnitude would be greater than the upper limit of the dynamic range of the display, the stimulus was rejected and replaced by a new one.

14. H. Sedgwick, "Space perception," in *Handbook of Perception and Human Performance*, K. Boff, L. Kaufman, and J. Thomas, eds. (Wiley, New York, 1986), Chap. 21.

15. J. M. Loomis and J. M. Knapp, "Visual perception of egocentric distance in real and virtual environments," in *Virtual and Adaptive Environments*, J. Hettinger and M. W. Haas, eds. (Erlbaum, Mahwah, N.J., 2003), pp. 21–46.

16. J. P. Frisby, D. Buckley, and P. A. Duke, "Evidence for good recovery of lengths of real objects seen with natural stereo viewing," Perception **25**, 129–154 (1996).

17. W. C. Gogel, "Equidistance tendency and its consequences," Psychol. Bull. **64**, 153–163 (1965).

18. G. Mather, "The use of image blur as a depth cue," Perception **26**, 1147–1158 (1997).

19. R. van Ee and C. J. Erkelens, "Temporal aspects of binocular slant perception," Vision Res. **36**, 43–51 (1996).

20. M. S. Banks and J. L. Dannemiller, "Visual psychophysics," in *Handbook of Infant Perception*, P. Salapatek and L. B. Cohen, eds. (Academic, New York, 1987), pp. 115–184.

21. D. Buckley, J. P. Frisby, and A. Blake, "Does the human visual system implement an ideal observer theory of slant from texture?" Vision Res. **36**, 1163–1176 (1996).

22. D. C. Knill, "Ideal observer perturbation analysis reveals human strategies for inferring surface orientation from texture," Vision Res. **38**, 2635–2656 (1998).

23. D. C. Knill and J. A. Saunders, "Do humans optimally integrate stereo and texture information for judgments of surface slant?" Vision Res. **43**, 2539–2558 (2003).

24. B. T. Backus and M. S. Banks, "Estimator reliability and distance scaling in stereoscopic slant perception," Perception **28**, 217–242 (1999).